

PHILOSOPHICAL SUBJECTS:  
ESSAYS PRESENTED TO P. F. STRAWSON

# PHILOSOPHICAL SUBJECTS

---

*Essays Presented to  
P. F. Strawson*

EDITED BY  
ZAK VAN STRAATEN

as Assistant Lecturer in  
lege of North Wales (1946-7).  
lectureship at University Col-  
of University College in 1948,  
ity years (1948-68). He was  
hen Gilbert Ryle retired in  
Waynslete Professor of  
tment to one of the four  
of Oxford meant that

CLARENDON PRESS  
1980

*Oxford University Press, Walton Street, Oxford OX2 6DP*

OXFORD LONDON GLASGOW

NEW YORK TORONTO MELBOURNE WELLINGTON

IBADAN NAIROBI DAR ES SALAAM LUSAKA CAPE TOWN

KUALA LUMPUR SINGAPORE JAKARTA HONG KONG TOKYO

DELHI BOMBAY CALCUTTA MADRAS KARACHI

© The several contributors listed on pages v and vi, and, so far as the residual copyright in the whole work is concerned, Oxford University Press 1990

*All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of Oxford University Press*

Published in the United States by  
Oxford University Press, New York

**British Library Cataloguing in Publication Data**

Philosophical subjects.

1. Philosophy—Addresses, essays, lectures

I. Strawson, Peter Frederick

II. Van Straaten, Zak

100 B29 80-40185

ISBN 0-19-824603-X

## Preface

Professor Strawson is a philosopher's philosopher. The philosophical literature of the last twenty-five years exhibits the considerable influence which his work has had on professional philosophers. The wealth of ideas which occur in his works in such fields as the philosophy of logic, logic, the philosophy of language, the philosophy of mind, Kant scholarship, aesthetics, and moral philosophy have produced a lively secondary literature. Two prominent instances are his debates with Russell on the theory of descriptions, and with Quine on aspects of philosophical logic.

In the present volume twelve philosophers present essays to P. F. Strawson on the occasion of his sixtieth birthday. All of the essays are original and occur here for the first time. The great variety of topics canvassed in the essays reflects the broad sweep of Professor Strawson's philosophical interests.

Professor Strawson has replied to all the contributors, and it is his replies, taken as a whole, which give a valuable and unique insight into his philosophical beliefs.

Peter Frederick Strawson was born on 23 November 1919. He was educated at Christ's College, Finchley, and St. John's College, Oxford, where he read Philosophy, Politics, and Economics, graduating with a BA in 1940. His career was interrupted by war service (1940-6) in the Royal Artillery, and the Royal Electrical and Mechanical Engineers. Professor Strawson's first teaching position was as Assistant Lecturer in Philosophy at the University College of North Wales (1946-7). In 1947 he was appointed to a Lectureship at University College, Oxford, becoming a fellow of University College in 1948, where he remained for the next twenty years (1948-68). He was promoted to Reader in 1966 and when Gilbert Ryle retired in 1968 Strawson succeeded him as Waynflete Professor of Metaphysical Philosophy. His appointment to one of the four chairs of philosophy in the University of Oxford meant that

not employ an unexplained 'could have', but only the relatively plain 'would have' which occurs in the form 'If he had thought ..., he would have decided ...'. Furthermore, they describe accountability in a way which explains why the concept's limits lie where they do. Without that explanatory component, Schlickian theories would not be seriously interesting.

6. WHY SCHLICKIAN THEORIES ARE UNACCEPTABLE

With it, however, they are in trouble: the Schlickian description of what accountability is—or of what the concept is *for*—strikes everyone as incomplete and strikes most people as wrong. The latter will say that although a distinction based on the utility of a certain sort of therapy, or behaviour-control might *coincide* with accountability/non-accountability, it cannot give the latter's essence, and that the Schlickian account of what the line is *for* does nothing like justice to the real nature of our praise- and blame-related responses. When we express indignation for someone's cruelty, or admiration for his unselfishness, we usually are not engaged in any sort of therapy: blame-related responses all involve something like hostility towards the subject, whereas a moral-pressure therapist, though he may have to feign ill-feeling for therapeutic purposes, can in fact be in a perfectly sunlit frame of mind. And—to move briefly to the 'welcome' side of the fence—one may apply moral pressures to encourage a welcomed kind of behaviour while remaining in an ice-cold frame of mind, with no feelings of gratitude, admiration of the like.

Schlickians defend their omission, arguing that we ought to jettison 'blame-related responses and handle ill-doers purely with a view to producing the best possible outcome. But what about the praise-related responses? Schlickians never say that we should give up admiration and gratitude and settle for 'therapies' aimed at encouraging recurrences of the welcomed kind of behaviour; but shouldn't they say just that? If blame-related responses are condemned just because we cannot explain the extent of accountability except by relating it to the relevance of a certain kind of 'therapy', then there is a strictly analogous case against the praise-related responses; but obviously we ought not to give up admiration and gratitude. So something has gone wrong.

We need to make room for at least part of what Schlickian theories omit, doing so in a manner which is not embarrassed by renewed difficulties over explaining why the line falls where it does. This double need is, in my view, satisfied by 'Freedom and Resentment'.

## 7. REACTIVE FEELINGS

According to Strawson, all that is omitted by Schlickian theories is the element of what he calls *reactive attitudes*, from which I shall at first lift out the component notion of a *reactive feeling*. Reactive feelings are ones which are prominent in blame, reproach, vilification, resentment, admiration, gratitude, praise, and so on. (If I could define 'reactive' I would do so, rather than resorting to examples. The problem of definition will be discussed in §§13-15, below.) Clearly, Schlickian theories offer us a way of handling accountability, or some notion coextensive with it, in a manner which does not demand reactive feelings. It is a manner which does demand the objective attitude towards the person concerned. The phrase 'the objective attitude' is Strawson's, and the core of its meaning seems to be this: To adopt the objective attitude towards something is to inquire into how it is structured and/or how it functions. The notion of objectivity will be further discussed in §15, below.

Many people find that feelings such as those of resentment and gratitude, indignation and admiration, do not easily occupy the mind along with a thoroughgoing concern to study the subject's behaviour patterns. Strawson mentions one familiar manifestation of this conflict between reactive feelings and objectivity of attitude—namely the common ability to dispel a hostile reactive feeling by cultivating objectivity of attitude towards the offender, e.g. dispelling indignation by viewing him as 'a case'.

In so far as reactive feelings won't mix with thorough objectivity, to that extent we must choose: we cannot always proceed as Schlick would have us do and throw in reactivity for good measure. Now, really, all that Schlickian theory advocates is that we let our response to each welcomed or regretted action be guided by a concern for achieving the best over-all outcome. So if reactive feelings are to have a place in our lives, we cannot

offers something which could explain the power of Spinozism to create doubts about accountability. It is that one element in 'moral ideas influenced by Christianity' is the thought 'that moral worth must be separated from any natural advantage whatsoever', a thought which led Kant 'to the conclusion that the source of moral thought and action must be located outside the empirically conditioned self'.

Anyone who thinks that moral worth cannot depend upon natural advantage is bound to see Spinozism as ruling out accountability. For Spinozism shows how all the facts about a man can be viewed as natural (dis)advantages which he has. (Even his existence. My very existence could be my bad luck, although there is no separate I who *has* the bad luck to *have* the existence.) To see this, consider that under certain genetic circumstances a person's very existence could be regarded as an injury or wrong done to him by his parents, although there is no separable *he* on whom the existence, and thus the injury, were inflicted.) I am sure that the Christian-Kantian thought does explain why Spinozism is thought to abolish accountability. And in representing the source of the phenomenon as positive, unitary, and deep-rooted, it is much truer to the observed facts than is the explanation which traces it to a scatter of conceptual failures.

I am not defending the Christian-Kantian thought. On the contrary, I agree with Williams: 'No human characteristic which is relevant to degrees of moral esteem can escape being an empirical characteristic, subject to empirical conditions, psychological history, and individual variation'; which is why Kant's treatment of moral worth, moral desert, accountability etc. is 'a shattering failure'. In fact, the Christian-Kantian thought generates a logically unsatisfiable concept of accountability—one which will not let a person count as blameworthy unless he is 'free' or 'autonomous' or 'self-made' in a sense in which it is logically impossible to be free or autonomous or self-made.

mon thinking about accountability. How widespread is it throughout the populace? In the minds of those who have it, is it sufficiently dominant to be constitutive of their 'concept of accountability'? I suspect that it is very widespread, but not very dominant; so I hesitate to say that the 'the ordinary concept of accountability' is 'logically unsatisfiable'. On the other hand, Strawson's claim to have provided for 'all we mean when ... we speak of desert' (p. 23), and his blunt dismissal of the sceptical view that 'blame is metaphysical' (p. 24), do, perhaps underestimate the strength of the Christian-Kantian thought. In particular, they do not do justice to the fact that although the plain man's moral thinking is not dominated by that thought, it is not informed either by any coherent alternative to it. But the point is not worth pursuing. We all know what happens when ordinary thoughtful people are pressed to get clear about accountability; and it doesn't matter how, if at all, the facts can be expressed as a thesis about 'the ordinary concept of accountability'.

What is interesting is the question of why the Christian-Kantian thought should have any acceptance at all. The explanation must lie partly in Christianity itself. If God is ultimately responsible for every fact about the natural realm, and is also the arbiter and punisher of wrongdoing, then there is something perfectly repellent about the idea of someone's being blameworthy for an action which is a truly natural event, whether fully caused or not. The God of Christianity cannot justly blame us for anything, it seems; unless he has given us some ultimate kind of agency which takes our actions right out of his field of operations, so that they cannot be fully explained by any combination of natural causes and of random events of the sort that would count as disorderly conduct on God's part. Thus, Christianity puts pressure upon its adherents to base accountability on some ultimate, not-of-this-world kind of autonomy, a kind which I, with Williams and others, believe to be logically impossible.

But the Christian-Kantian thought extends more widely than does any considered theology of God as creator and judge; and although that may be partly due to inertia, there is also a different explanation for the Christian-Kantian thought, which I now present. It is mine, I think, rather than Strawson's;



my endeavour to learn how to get him to repeat a certain kind of behaviour cannot be explained by my gratitude to him. Even if I am grateful to him, what explains my teleological inquiry must be something else—presumably my desire for a repetition of the welcomed kind of behaviour. Or so I maintain.

I now add a further suggestion. The attitude of  $x$  towards  $y$  is reactive *if and only if* it is a pro or con attitude which could not explain  $x$ 's engaging in teleological inquiry into how  $y$  works. I cannot find any convincing counter-examples to this biconditional.

It does not make 'reactive' and 'objective' complementary within the class of pro and con attitudes. What I propose is this: if  $x$ 's attitude to  $y$  *does* generate teleological inquiry, to that extent it is objective; if it *could not* do so, it is reactive; it is *could but does not*, it is neither.

Consider the attitude of Dives towards Lazarus when he passes him on the street and feels fastidious disgust at his stink, his sores, his cringing whine. He hurries on: so far from inquiring, he tries to shut out the facts by shutting out the sensory intake; and so it would be an abuse of language to call his attitude towards Lazarus 'objective'. But it does not belong with the attitudes which Strawson calls 'reactive': it is too far removed from resentment, indignation, and blame, and is too close to what Lazarus might feel about a dead animal in the gutter. These intuitions about Dives' disgust are nicely confirmed by my suggested criteria. The disgust is not reactive because it *could* explain teleological inquiries: if Dives studied Lazarus's condition with a view to ameliorating it or having him arrested, etc. this could be because he found Lazarus disgusting. But this episode of disgust is not objective because it does not, in fact, lead Dives to teleological inquiries.

However, my proposed necessary and sufficient condition for an attitude to be 'reactive' is not useful or satisfying. It says only that reactivity is incompatible with a certain causal or explanatory power—i.e. that  $x$ 's attitude cannot both be reactive and be the cause or explanation of  $x$ 's engaging in teleological inquiry. This doesn't tell us what reactivity is, but only names one of its properties—and a negative, modal, relational property to boot.

So I have failed to say in general terms what

or what the kind of interpersonal relationship is which defines the home ground of personal reactive attitudes. But I still think that Strawson has made a profound contribution to the understanding of accountability; for his account could do its principal work without using 'reactive' or any substitute for it, and merely speaking separately of 'resentment', 'gratitude', 'indignation' and a few others.

#### 16. THE OBJECTIVE-REACTIVE CONFLICT

Still, it is regrettable that we still have no satisfactory theoretic unification of these matters. If we had one, it might help us to understand the tendency of reactive attitudes to be in conflict or in tension with thorough objectivity. What is the source of this conflict or tension, and how strong is it?

As regards the latter question: it is trivially true that an *exclusively* objective frame of mind must absolutely exclude everything non-objective; but the question is whether a *thoroughly* objective frame of mind can be combined with reactive attitudes. Not in me, it seems; and many of my acquaintances say the same. So would Dostoyevsky:

But what can I do if I don't even feel resentment? . . . My anger, in consequence of the damned laws of consciousness, is subject to chemical decomposition. As you look, its object vanishes into thin air, its reasons evaporate, the offender is nowhere to be found, the affront ceases to be an offence and becomes destiny, something like tooth-ache, for which nobody is to blame.<sup>13</sup>

Perhaps there is a 'law of consciousness' by which thorough objectivity casts out reactive attitudes. Some people deny that this dependably happens when *they* 'look' with Dostoyevsky or 'measure the lot' with Yeats; but this conflict of testimony could reflect our unclarity as to what the issue is. On the other hand, the phenomenon may be subject to real interpersonal variation.

Strawson seems undecided about the strength of the conflict. He makes it absolute when he says that the objective attitude 'cannot' be emotionally toned with resentment, gratitude, etc. (p. 9), and when he alludes to the 'human isolation' which 'a sustained objectivity of interpersonal attitude . . . would entail' (p. 11). But his reference to 'the tension there is, in us'

<sup>13</sup> F. Dostoyevsky, *Notes from Underground* (Penguin Books, 1972, trans. J. Goulson), p. 27 (Ch. 1, §5).

show goodwill towards James, then you cannot have acknowledged James's *claim* to goodwill from John or endorsed his *demand* for it.' The words 'claim' and 'demand' occur frequently in Strawson's discussion; and I concede that it is plausible to link 'demand for x' with the idea of indignation at x's not being given. But this link cannot be used to explain what 'indignation' is. If it has any explanatory value, it runs the other way, enabling us to explain 'demand' by reference to indignation, as Strawson does: 'The making of the demand is the proneness to such attitudes' as those of 'disapprobation and indignation' (p. 22). Anyway, I doubt if 'demand' really covers all the ground: I can find no place for it in describing such undisappointed reactive feelings as those of gratitude and reciprocating love.

Perhaps we should try to fill out the treatment of impersonal and self-directed reactive attitudes by restoring the interpersonal-relation element. But Strawson doesn't want us to do this, and one can see why: if my indignation at your treatment of him is essentially to involve an interpersonal relationship, who are its relata—you and I? or you and he? Each answer seems peculiar. Yet I wish to defend the answer 'You and I'; but first I must return to the home ground of *personal* reactive attitudes, for Strawson's treatment of them contains a small defect whose consequences are now plaguing us.

## 18. RELOCATING ONE ELEMENT

Personal reactive attitudes are introduced, as I have noted, through the notions of claim-to-goodwill and interpersonal relation. Now, Strawson speaks of the goodwill which is demanded in certain relations; but isn't it also demanded outside of them? And cannot the latter demands also generate reactive attitudes? And, to take in some of the territory not covered by 'demand' and 'claim', cannot gratitude, for instance, occur without any antecedent personal involvement? But in all these cases there is the "involvement" created by the very behaviour to which the reactive attitude is a response. That is true, but I am sure that Strawson is not thinking of kinds of involvement or 'interpersonal relation' which could be created just by kicking somebody or throwing him a coin. A large theme in 'Freedom and Resentment' is the contrast between the involvements which go with reactive attitudes and

the 'isolation' which would be entailed by their absence, as well as the 'relief from the strains of involvement' (p. 12) which can be brought about by a temporary replacement of reactive by objective attitudes; and all of that is reduced to nonsense if one construes 'involvement' etc. so as to include mere helpings and harmings.

As for the converse: Strawson himself clearly implies that the fact that ill-will occurs within a relationship of the emphasized sort does not guarantee that the response to it will be reactive.

One might conclude that the notion of interpersonal relation is not supposed to help explain what a personal reactive attitude is, and is offered only as part of the natural history of reactive attitudes—a mere description of their place in our lives. But that is hard to reconcile with the amount of weight Strawson seems to lay upon such expressions as 'participant' and 'non-detached'. Fortunately, there is another way out.

It is to give the notion of an interpersonal relation (of the relevant kind) a role in the analytic or explanatory part of the account, but not quite the role initially allotted to it by Strawson. What should be emphasized, I suggest, is not the relations *within which* reactive attitudes arise, but rather the relations *towards which* they point. If I resent someone's treatment of me, there may have been antecedently no special kind of relation between us; but my very resentment creates one, or sets the stage for one. I cannot say what the 'special kind' is: that is the problem which I could not solve when trying to define 'reactive' in §14 above. But still I submit that it helps if those participations and involvements which Strawson emphasizes are seen not primarily as the ground in which reactive attitudes grow, but rather as embodied in or consequential upon them; not as required in the past or present, but as implied or suggested or invited for the future.

This idea is implicit throughout most of 'Freedom and Resentment', which is why I could quietly adopt it in my section on interpersonal relations. But it needs to be made more emphatic and explicit than Strawson makes it. It could lead to a tightening of the curiously loose and structureless paragraph in which 'reactive attitudes' are first introduced (p. 6). It could also let us strengthen a soft spot in Strawson's rationale for the

Essentially the same point holds for non-moral reactive attitudes; but here there is a terminological snag. Strawson assumes—rightly, in my view—that an attitude counts as ‘moral’ only if it rests on a general principle, or anyway on something which does not essentially refer to any particular item. So an attitude of mine is not moral if its basis essentially involves myself: I am *morally indignant* at your contemptuous attitude towards a benefactor, but I *resent* your contemptuous attitude towards *me*. But the basis for an attitude might lack generality—thus depriving the attitude of the status of ‘moral’—in some quite different way. For a bit of behaviour might enrage me on a particular occasion, although it neither infringes any general principle which I hold nor essentially involves myself. For instance, I take no general stand on attitudes to natural beauty, but on this one occasion it just makes me angry to see a man walk unheeding past the masses of Alpine Lilies and Indian Paintbrush. Or I have an unreasoned ‘thing’ about Bruckner, which leads me to feel something like gratitude towards anyone who loves his music. That anger and that ‘gratitude’ are both reactive, I suppose; but they are not ‘moral’, since one concerns a particular occasion and the other a particular person, and neither rests on general principles. But neither of them fits comfortably under Strawson’s label ‘*personal* reactive attitude’, since that label so naturally suggests an attitude which responds to someone’s attitude to *oneself*.

I suggest, therefore, that the two basic categories of reactive attitudes are ‘non-principled’ and ‘principled’ (or ‘moral’), with ‘personal’ as an important species within the former.

As for self-reactive attitudes: some of them are principled and some are not; for an attitude of self-censure or self-congratulation may, but need not, rest upon some principle which one holds. Strawson focuses primarily on the principled ones—which rest on one’s acknowledgement of others’ claims on one’s goodwill—and perhaps they matter most. But there are others, such as self-reproach for having made a fool of oneself in public. Incidentally, Strawson’s use of the word ‘moral’ is unsatisfactory on any showing, for his ‘moral’ category positively excludes self-reactive attitudes. Admittedly, what is ‘moral’ must have a general basis; but that is no

obstacle to allowing that self-reactive attitudes can be moral. My remorse over my cruelty is as principled as my indignation over yours.

So much for taxonomy and terminology. Returning now to the substantive point: I suggest that although it is all right to tie reactive attitudes to responses to *somebody's attitude*, it is unduly narrowing to tie them to responses to *somebody's attitude towards somebody*. The '... towards somebody' bit must loom large in the natural history—and perhaps also in the ethics—of reactive attitudes, but not in the account of what reactivity is.

When the account is thus generalized, it covers cases which are untouched by Strawson's treatment. Also, as I have shown, it forces us to make independently worthwhile revisions in the taxonomy and terminology. And, finally, it makes no difference to the relevant part of Strawson's rationale for the line around accountability. For that concerns cases where, despite appearances to the contrary, 'the agent's attitude and intentions [are] what we demand they should be' (p. 7); and this need not draw on the idea that the relevant 'attitude' must be towards one or more people.<sup>14</sup>

<sup>14</sup> An earlier version of this paper was the subject of extremely helpful comments by Peter Remnant, Richard Sikora, Earl Winkler, and, especially, Donald G. Brown.

and 'I see a variety of colours'. It is a corollary of her view that 'No problem of the continuity or reidentification of "the I" can arise' (p. 62); it is merely that the same human being has different I-thoughts at different times. A 'lapse of self-consciousness' is just a lack of 'unmediated agent-or-patient conceptions of actions, happenings, and states'. But these conceptions, she maintains, 'do not involve the connection of what is understood by a predicate with a distinctly conceived subject'. The notion of such a subject is just a '(deeply rooted) grammatical illusion' (p. 65).

Strawson, too, speaks of an illusion, though not a grammatical one. He argues that Kant exposed and, less completely, explained 'a natural and powerful illusion' by which 'we mistake the necessary unity of consciousness for ... an awareness of a unitary subject' (p. 163). Reconstructing and completing Kant's account, he says that while 'a certain character of connectedness and unity' in a temporally extended series of experiences, yielding 'experience of a unified objective world', 'makes room for the idea of one subjective or experiential route through the world', and so provides 'as it were, the basic ground for the possibility of an empirical use for the concept of the subject of such an autobiography, the concept of the self', this is not in-itself enough. It is also vital that 'our ordinary concept of personal identity does carry with it empirically applicable criteria for the numerical identity through time of a subject of experiences (a man or human being) ... which involve an essential reference to the human body'. This 'supplies an absolutely firm basis for a genuinely object-referring use of names and personal pronouns, in sentences in which states of consciousness, inner experiences, are ascribed to the objects referred to by the names or pronouns' (pp. 163-4). But 'the fact that lies at the root of the Cartesian illusion' is that 'When a man (a subject of experience) ascribes a current or directly remembered state of consciousness to himself, no use whatever of any criteria of personal identity is required to justify his use of the pronoun "I" to refer to the subject of that experience' (pp. 164-5). Yet Strawson insists that when 'I' is thus used in 'criterionless self-ascription', it does not 'lose its role of referring to a subject'. He suggests that it can still refer to a subject because either it issues from the mouth of an empiri-

cally identifiable person or, if used in soliloquy, it is used by a person who would acknowledge the applicability of ordinary empirical criteria to the question of his own identity with someone who performed some earlier action. "I" can be used without criteria of subject-identity and yet refer to a subject, because, even in such a use, the links with those criteria are not in practice severed" (p. 165).

Strawson, then, is partly in agreement but partly in disagreement with Anscombe and Vendler. Like them he holds that 'a purely inner and yet subject-referring use for "I"' is an illusion. Consequently 'If we try to abstract this use, to shake off the connection with ordinary criteria of personal identity ... what we really do is simply to deprive our use of "I" of any referential force whatever. It will simply express, as Kant would say, "consciousness in general"' (p. 166). But what Strawson here says that we may try to do, Anscombe and Vendler think we have done. They say that the central use of 'I' is this non-referring one, whereas Strawson says that the ordinary use remains referential, because the links with empirical criteria of identity are not in practice severed. (Also, unlike Anscombe, Strawson does not shy away from the characteristically Cartesian thoughts: for him it is especially states of consciousness, inner experiences, that are self-ascribed immediately and without criteria.)

We have, then, two issues. Are Strawson, Anscombe, and Vendler all in the right in their general support of a Kantian against a Cartesian view? And, where they differ, is Strawson right against the others in saying that 'I' remains a referring term?

On this second question, the view of Anscombe and Vendler is paradoxical. Surely ordinary speakers commonly at least intend to use 'I' as a referring term. On first learning the jargon they would say confidently that it is one. If this is a grammatical illusion, it is one nearly all of us share. Also, as Anscombe admits, the rule for the truth of propositions with 'I' as subject is

If *X* makes assertions with 'I' as subject, then those assertions will be true if and only if the predicates thus



become disembodied, which is conceivable though it may well be physically impossible, and it is abandoned when, with Vendler, we engage in feats of transference. And in these marginal cases the over-determination of meaning by the two rules does cause trouble: in employing one rule we have to ignore the other.

This brings us to Vendler's argument, which seems straightforward and, given its assumptions, conclusive. It is necessary that Z.V. is not Claudius. But I can imagine my being Claudius, so it is not necessary that I am not Claudius. From this it does not, indeed, follow that I am not Z.V. But it does follow that I am not necessarily Z.V., and hence that 'I am Z.V.' cannot be a true identity statement if both 'I' and 'Z.V.' are rigid designators. 'I' cannot be used as a rigid designator, as a name or a definite description used referentially, both in 'I am Z.V.' and in 'I might have been Claudius'.

Once it is thus set out explicitly, we can see that the argument does not show that 'I' is never used to refer to a thing (or person or human being), but only that there are some uses in which it does not rigidly designate a human being, and hence that if it does then rigidly designate anything, this will have to be a Cartesian Ego. It follows that if there is another use in which it does rigidly designate the human being who uses it, 'I' must be ambiguous. But this is only to put in other words what was suggested above, that 'I' is governed by two meaning rules which initially diverge, though they converge again in ordinary circumstances. Vendler's examples of transference stress that there are special kinds of thinking in which the divergence is unresolved. I as a subject of consciousness, or perhaps as a mere pseudo-subject in a series of co-conscious experiences, might have been Claudius, though I, this human being, could not.

But should we speak here of a subject or of a pseudo-subject? Vendler is wrong to say that the 'I' which is the subject of transference is 'the bare form of consciousness', though he is right to say that it is a frame in which any picture fits. When I imagine being at Cannae, it is *my* being at Cannae that I imagine, not that of a bare form of consciousness. (After all, a good many thousand 'forms of consciousness' were there, but this is not the central part of what I imagine.) Yet it is not, in general, that I imagine there being experiences linked by co-

consciousness to my actual ones: I do not have to think of myself, at Cannae, remembering the twentieth century. I ascribe the imagined experiences to a subject. I need not imagine them as linked by co-consciousness to the present series, and *a fortiori* I do not merely do this. (Nor is it merely that I am joining them to this series just by imagining them: for that would be true of anything I imagined, and so cannot explain the distinctive element of my imagining *my* being there.) But still we have the nagging question, is it to a real subject or to a pseudo-subject that I ascribe them? As we noted, there is a real subject of my actual experiences—these imaginings among them—namely this human being. So when I imagine whatever is the subject of my actual experiences as being at Cannae, it is to this human being that I in effect ascribe the imagined experiences; but not directly, not explicitly, and that is how I avoid the contradiction of supposing this human being to have lived at a different time.

The two premisses of Vendler's main argument—'Z.V. could not have lived at a different period' and 'I could have lived at a different period'—both call for fuller interpretation. As I have argued elsewhere about such Kripkean modal theses, the first really means only that we have a way of thinking and speaking about the same individual with respect to counterfactual possibilities that precludes this. We think of the same individual, even in counterfactual possibilities, as being whatever persistent thing of that kind has the same origin.<sup>2</sup> Possibilities for this individual (human or non-human) are possible divergences from its actual history. But the second premiss brings out another way in which I can think, with respect to counterfactual possibilities, of this special individual, myself. I hold on to myself merely as the subject, whatever it may be, of these present experiences. The counterfactually imagined experiences are made its simply by direct fiat. I do not construct an apparently objective course of events and then have to identify myself with some item already in it: rather I tell a story about myself as centre.

It may be objected that if the subject of these present experiences happens to be this human being, then it could not

<sup>2</sup> Cf my 'De What Re is De Re Modality?', in *The Journal of Philosophy*, Vol 71 (1974), pp 551-61, and my *Problems from Locke* (Oxford, 1976), pp 152-9.

Strawson also says: 'It is not easy to become intensely aware of the immediate character, of the purely inner basis, of such self-ascription while both retaining the sense of ascription to a subject and forgetting that immediate reports of experience have this character of ascriptions to a subject only because of the links I have mentioned with ordinary criteria of personal identity' (p. 166). What he speaks of forgetting is what I would deny. Since, as he says, such self-ascription is immediate and has a purely inner basis—what I have called being joined on to a series of co-conscious experiences—what *present* role can the alleged links with ordinary criteria play? Strawson does not explain this, and it is hard to see how they can be playing any role, in view of ways in which they can be broken or dispensed with, for example in thought about mediums and possession or in a Butler–Swinburne–everyday concept of the absolute identity of a person.

On the other hand, taking up a theme that was set aside earlier, we can allow that these empirical criteria are historically important: they will have contributed to the development of the concept of a unitary subject to which experiences (etc.) are ascribed. In what he calls 'the coup de grâce to Cartesianism', Strawson, following Kant, argues that Descartes's rational psychology is powerless to defend its conclusions against alternative but equally empty theories. Perhaps there are a thousand souls simultaneously thinking the thoughts that one man's utterance expresses. 'How could the—or each—soul persuade itself of its uniqueness?' (p. 168). Anscombe echoes this: 'How do I know that "I" is not ten thinkers thinking in unison? Or perhaps not quite succeeding. That might account for the confusion of thought which I sometimes feel' (p. 53). But these intendedly rhetorical questions themselves hint at an answer. The approximation to harmony that my thoughts display is some evidence, though not conclusive evidence, of a single thinker, of a source whose parts, if it has parts, are closely interconnected. As jam sessions go, my mental performances are fairly successful. If there were a thousand, or even ten, souls thinking in near unison, this would call for some explanation of how it was achieved. Nevertheless, it is obvious that historically the notion of a single subject of experiences is borrowed from that of a single human being: in particular, it is other people's

view of me as one person that has consolidated my unitary view of myself. Again, it is no accident that what I have called the two senses of 'I' converge. They grew up together: they developed in virtue of their convergence in the ordinary central uses. Each of us learned to use 'I' to refer to the subject whatever it may be, of his co-conscious experiences because he was having those experiences in circumstances in which the corresponding predicates were true, and many of them observably true, of this human being. So when Strawson says that 'immediate reports of experience have this character of ascriptions to a subject only because of the links . . . with ordinary criteria of personal identity', we can accept this if 'because' has a causal, historical, sense; what I deny is that we should accept it if 'because' stands for a contemporaneous linguistic constraint.

We finally return, then, to the other major issue, that between Kant and Descartes. Am I now rejecting what Strawson takes as Kant's insight and opening the way for the acceptance of a Cartesian Ego after all? Not quite. I am conceding to Descartes that whatever sophistication and learning may lie behind our present competence, we do now have a way of ascribing experiences, states, and actions to a supposed single subject, independently of any criteria for its unity and identity through time and without any contemporaneous reliance on the equation of this subject with a body or a human being, and we can therefore coherently engage in feats of transference or speculations about disembodiment which would conflict with such an equation. My essence is thinking and nothing but thinking in the sense that thinking alone is contemporaneously involved in this ascription. But I agree with Kant and Strawson and Anscombe and Vendler, against Descartes, that such a method of ascription does not suffice to introduce a thing to which these items belong and whose existence involves thinking alone, which is ontologically distinct from the body and of which we know, just in virtue of this method of ascription, that this thing can continue to exist whether there is a body or not. Descartes's mistake was to suppose that something whose style of introduction into our thought and speech involved nothing but thinking was thereby shown to be such that its existence involved nothing but thinking. The truth of the matter is that 'I'

# The Primitiveness of the Concept of a Person<sup>1</sup>

Hidé Ishiguro

1. Professor Strawson said in his book *Individuals* that the concept of a person was primitive. He made this claim in order to show up certain conceptual difficulties involved in two different theories which still leave their traces in philosophy. One is the Cartesian view, and the other is what Strawson calls the no-ownership view of experiences. But Strawson's positive view of the concept of a person has been criticized even by those who neither hold Cartesian views nor are tempted by any no-ownership view of experiences. For example, it has been suggested that what Strawson himself says about m-predicates (material predicates) and p-predicates (person predicates) entails that the concept of a person is not primitive.

I think that there is something very important in the claim that the concept of a person is primitive. Strawson's claims will enable us to develop a theory which tells us why the question whether certain beings are persons and others are not admits of an objective answer (rather than an answer based on our arbitrary decision). And as the answer does not involve reference to other kinds of entity besides persons, it does not require us to believe in the existence of souls as entities within us. Nor do we need to have a way of distinguishing persons in terms of purely physical characteristics.

Suppose, for example, that the concept of a person were like that of a friend. The application of 'friend' is based on a mutual decision to treat one another in a special way. If 'person' were similar, one would be mistaken about who was to count as persons only if one was mistaken about their attitude towards oneself. (Just as one may be mistaken about another being a friend if the other is actually quite indifferent to one.) In this case one could never say that it was wrong for someone not to

regard blacks as persons, or that a particular society was mistaken in not treating madmen as persons. But we want to say this. It seems not to depend on our personal decision or even group decisions that some beings are persons—that they belong to a particular sort of thing to which we belong. It is something which we can make mistakes about and come to discover that we have done so. (The fact that there are borderline cases and the fact that differences between kinds of things are continuous do not tell against their being questions with objective answers.) Nevertheless I want to argue that; when one realizes that one was mistaken in not treating a being as a person, the basis of the realization is very different in kind from the basis upon which we come to realize, for example, that we were mistaken in treating whales as fish and not mammals. It is not like discovering a hitherto unnoticed morphological feature which the being shares with other human beings (as it was in the case of whales).

Certain points made in the last decade by Hilary Putnam and Saul Kripke and others on the relation of reference and meaning of certain words, will illuminate in an unexpected way what was so interesting in Strawson's claim. But Strawson's view of persons will actually take us further than have recent arguments about the referential theory of meaning. It suggests that such supposed dilemmas as 'Does one understand what a concept is by indexical means or by giving a defining property?' are just as misleading as was Russell's division of knowledge into knowledge by acquaintance and knowledge by description. These are not exhaustive alternatives. How one comes to understand what sort of things fall under a concept usually involves much more than either pointing, defining, or describing.

I will not at this point try to specify completely what is meant by a concept being primitive in the interesting sense Strawson requires. I hope it will become clearer during this paper. Nevertheless I will begin by saying a few things about it. First, to say that a concept is primitive is not to say that it is simple. Many complex concepts are primitive in a sense we will explain later. We shall see that primitiveness has more to do with the indispensability of a concept than with its simplicity. In the second place, primitiveness is not a feature of concepts in any

clear and indistinct (confused) at the same time. Clarity does not always come together with distinctness. An idea is clear if we can clearly distinguish from others, things that fall under the idea. Yet the same idea may be indistinct if we do not understand the complex structure of the idea, i.e. if we do not know the properties that all and only these things which fall under the concept have. With ideas of abstract things like numbers it is hardly possible to have a clear idea without having a distinct idea. For example we cannot distinguish odd numbers from even numbers, without knowing what are the properties that are necessary and sufficient to make a number odd, but for most other kinds of things we have ideas that are not distinct. Thus Leibniz says 'we cannot give a definition of them. We make them known only by examples' (N. E. II, 29, 4). The position is not simple as this remark of Leibniz might indicate, however. Even for material stuff concepts like 'gold' it is highly unlikely that we should have a clear idea unless we agreed on, or saw the prospect of agreeing on, a common distinct idea. This is because pointing to a sample does not by itself tell us how we decide that other objects are to count as being made of the same stuff as that sample. We might have been able so far to tell gold apart from other metals because every piece of gold we have encountered has been yellow and heavier than other metals. But if I encounter another alloy or another compound or even hitherto unknown metal which is of the same colour but heavier, how would I know whether it is to be considered the same metal or not? Without a knowledge of the relevant theory which would help me decide — i.e. without a distinct idea I would fail to distinguish gold from other metals. It is not enough to say that it is 'whatever bears a certain equivalence relation to the piece referred to as "this"'. We need to specify which equivalence relation it is or what it is to count as being of the same metal. Is an isotope of gold to count as the same kind of metal as this piece of gold? It is our theory about the sample in question which gives us the specific equivalence relation. So in many cases our clear ideas are also distinct ideas. cases our clear ideas are also distinct ideas.

When do we have clear but indistinct ideas? According to Leibniz, the best examples are ideas which correspond to Locke's secondary qualities. These ideas were simple ideas for

Locke, but for Leibniz these are complex ideas of complex properties. They are clear because we can distinguish things that fall under the idea *red* from other colours or hot things from cold things. They are nevertheless indistinct ideas because we do not understand the nature of the property of a thing's being red (e.g. that it reflects light waves of longest wave-length), or being hot (e.g. increased kinetic energy of the molecules). Those of us who have sight or sensation use our sight or touch to get things that are red or hot, rather than using a set of articulate criteria. Because our senses enable us to engage in discriminatory activities and a whole complex of activities dependent on these, we have a clear grasp of these properties which we cannot verbalize, having no grasp of the structure of things responsible for this property.

If we realize that there is a light wave with a wave length longer than that of light we hitherto called red light, we do not freely decide whether to think of this newly discovered light as red light. Red is the colour corresponding to certain perceptual experiences, and even if it was one of our beliefs that the colour was related to light of the longest wavelength, we *can* discover that our beliefs were wrong. Our senses, our perceptual experiences had more or less fixed the use of the term 'red', i.e. the extension of the concept.

5. I believe that it is similar in the case of the concept of a person. We do not recognize a person just by sight or by touch. But just as in the case of ideas of secondary qualities, we are hard put to define what a person is although we can more or less recognize one when we encounter one and see him act. By observing people we can discover various properties that all of them have. And even if on rare occasions we had to decide whether a being was a person on strength of its having or lacking certain property, there is normally no question of arbitrarily defining what goes into the concept of a person or what makes up its extension. We have some rough idea of what it is that we are committing ourselves to when we consider a being as a person. What matters to personhood of a being is the nature of the relations we find ourselves in, or decide to cultivate, or can even make ourselves see, however reluctantly, as being possible, to establish with the being. We are obliged to consider it as a person whether we like it or not. Our ability to



communicate mutually and to understand or enter into various relations gives the extension of the concept of a person to which we believe we belong. And this does not have to wait for the moment when we have identified them as material bodies having a certain set of *m*-properties. (Our ascribing *p*-predicates to them presupposes rather than explains our regarding them as persons.)

Now obviously the fact that we treat certain objects in certain ways, and believe that we have entered into certain relation with them, does not by itself determine what kind of object they are. A fetishist's worship of an inanimate object, and his talking to it or engaging in various ritualized acts with it, does not turn an inanimate object into a person. Nor does the mere fact that certain people talk to and believe they communicate with dogs, goldfish, or even plants, establish that dogs, goldfish or plants are persons. Minimally personhood has to involve the possibility of entering into *mutual* relationship of the above kind. Nevertheless, it appears that the extension of the concept 'person' is determined not so much by our grasp of a clear set of characteristics which apply to a subset of animals as by our ability to enter into certain relationships or imagine ourselves doing so and realize in this way that they are of a kind with ourselves.

Strawson has written interestingly about this feature of human life (which he calls 'shared reactive attitudes') in his essays on *Freedom and Resentment* (British Academy Lecture 1962). It has been my purpose to connect it with the idea that the concept of a person is primitive relative to the concept of a material body. I have attempted to establish the link in three successive steps. (a) We get the concept person by some understanding of *ourselves* and of beings, typical beings, with whom we can identify ourselves as being of a kind. (b) This means that we can distinguish persons from other things without being able to articulate what kind of material objects they are by neutral physical features. (c) This in turn shows that it is indispensable that we have the sortal concept of a person. And, as I have said, I take the primitiveness of a concept to mean the indispensability of one sortal concept relative to another sortal concept. Of course (b) and (c) are compatible with the fact that there are some structural physical feature common to all those beings we

consider as persons, as in the case of gold. But it does not entail this fact. What fall under the sortal *person* are not necessarily beings with the relevant genetic or physiological similarity to myself and standard persons—this may or may not turn out to be so—but beings who have the same range of dispositions, abilities, etc. Thus I would not like to commit myself to the view that a person is a natural kind. It may, for example, correspond to a disjunction of natural kinds, or one day include man-made beings which do not look like us. This will not affect the claim that the sortal concept is primitive.

People are sceptical about philosophical positions which start from *de facto* use of concepts. In the minds of many critics the fact that people can obviously disagree about concepts of misuse them adds to the contingency of their uses. Yet, just as no theory of number would be worth defending if it did not minimally give an account of our entrenched practice of counting, no theory of persons would be worth defending if it did not capture our practice of regarding one another as persons. The necessary link between number theory and the practice of counting and doing sums is not destroyed by the fact that we occasionally miscount and miscalculate, nor by the fact that there are people who cannot learn to count. Similarly the necessary connection which must obtain between any philosophical concept of person and our practice of using the sortal concept is not weakened by the existence of disagreements and puzzles about personhood. Professor Strawson has reminded us of this truth.

objectivity which rests upon interpersonal agreement. For the purpose of focusing upon a manageable section of our over-all conceptual scheme rather than because it represents a genuine conceptual possibility, Strawson pretends that his subject makes no allowance for the existence of other observers. (Throughout, then, 'existence unperceived' is, effectively, 'existence unperceived by me'.) The idea which *does* concern Strawson is the idea of an experience's being of something distinct from it, and therefore the idea of something which is capable of existing independently of any experience of it.

Now Strawson dubs his investigation the elaboration of the conditions for a 'non-solipsistic consciousness'. But there is implicit in his notion of objectivity a restriction which precludes the application of the title 'theory of the objective world' to views which their proponents did not believe were solipsistic. The proponents of these views have a yet more general notion of objectivity in mind, according to which a theory has objective significance if it comprises propositions which, if true at a given time, are not true in virtue of the state of the subject at that time. This differs from Strawson's notion, for there may be no question of the subject's *experiencing* the reality which is constituted by the truths that do not belong in his biography.

For example, the subject's experience may be regular in a way which allows him to express various conditional or counterfactual propositions about what would be experienced were such and such else to be experienced. These propositions, if true at all, are not true in virtue of any actual occurrence in the subject's biography, and may be taken<sup>2</sup> as the basis of a claim that an objective reality is thereby allowed for. However this may be, it is not an objective reality in the sense which concerns Strawson:

... I shall mean by a non-solipsistic consciousness, the consciousness of a being who has a use for the distinction between himself and his states on the one hand, and something not himself, or a state of himself, *of which he has experience* on the other.<sup>3</sup>

<sup>2</sup> They have been so taken by phenomenologists. See, for example, C. I. Lewis, *An Analysis of Knowledge and Valuation* (Open Court, La Salle, Illinois, 1946), pp. 226-30, and *Mind and the World Order* (Dover, New York, 1956), pp. 135-9.

<sup>3</sup> *Individuals* (Methuen, London, 1959), pp. 69: my italics.

For whatever it is that makes these purportedly 'objective' propositions true, it is not something of which the subject can be said to have experience.

## PART ONE

Intertwined in the chapter are several ideas on which a defence of the Kantian thesis might rest, but the main line of argument is one in which the need for space arises from the requirement that the subject of experience (hereafter 'Hero') be able to reidentify the objects of his experience. As Strawson summarizes the results of this chapter in a later one:

... we had to introduce, in auditory terms, an analogue of space in order to make room for the idea of reidentifiable particulars ...

A defence of the Kantian thesis on these lines must have two distinguishable stages; it has to be shown that the idea of reidentifiable objects is implicit in the idea of objectivity, and it has to be shown that criteria of reidentification, with the attendant distinction between qualitative and numerical identity, can only be framed in a spatial (or quasi-spatial) world.

On the first stage of the argument, Strawson has this to say:

... to have a conceptual scheme in which a distinction is made between oneself or one's states and auditory items which are not states of oneself, is to have a conceptual scheme in which the existence of auditory items is logically independent of the existence of one's states or of oneself. Thus it is to have a conceptual scheme in which it is logically possible that such items should exist whether or not they were being observed, and hence should continue to exist through an interval during which they were not being observed. So it seems that it must be the case that there could be reidentifiable particulars in a purely auditory world if the conditions of a non-solipsistic consciousness could be fulfilled for such a world. Now it might further be said that it makes no sense to say that there logically could be reidentifiable particulars in a purely auditory world, unless criteria for reidentification can be framed or devised in purely auditory terms. And if this is correct, as it seems to be, we have the conclusion that the conditions of a non-solipsistic consciousness

<sup>4</sup> *Individuals*, p. 118.

<sup>5</sup> I shall drop this qualification in what follows, except where the context indicates otherwise, I shall use 'spatial' in the weak sense of 'spatial or quasi-spatial'.

can be satisfied in such a world only if we can describe in purely auditory terms criteria for reidentification of sound particulars.<sup>6</sup>

The second stage of the argument is not filled out in any great detail, but perhaps it is obvious how it would run. Hero has to be able to distinguish among later experiences of qualitatively indistinguishable phenomena those that are, and those that are not, later stages of the same phenomenon that he experienced earlier. This can only be done by taking into account the *relations* in which the phenomena stand; more specifically, relations which do not hold in virtue of the intrinsic non-relational character of the things related. And perhaps this may be seen as an abstract formal description of spatial relations.

Ingenious though this argument is, there is room for considerable doubt about its cogency, especially when it is interpreted in such a way that Strawson's auditory world provides an illustration of it. In that world, qualitatively identical sounds may be distinguished by their 'location at' (by being heard with) different pitch-levels of the master-sound. Given that this is the case, one may well feel that no genuine distinction between qualitative and numerical identity has been provided. Since different pitch-levels of the master-sound are qualitatively distinguishable, auditory presentations of numerically distinct sounds are never qualitatively indistinguishable, so long as a sufficiently inclusive view is taken of what to apply this concept to. But in the present context, a weaker point will suffice. If Strawson's Hero uses a genuine criterion of numerical identity, then space is not necessary for him to do so. Since his criterion does not make any real use of the dimensionality provided by the continuous variation in the pitch-levels of the master-sound we may suppose with equal legitimacy that criteria of reidentification can be framed in an auditory universe in which an *unordered* series of master-sounds plays the same distinguishing role as the ordered series of pitch-levels of Strawson's single master-sound.<sup>7</sup>

<sup>6</sup> *Individuals*, pp. 72-3.

<sup>7</sup> I assume here and throughout that the existence of an intrinsic ordering between 'places' is essential to a space or a quasi-space. This certainly seems to be the framework within which Strawson was working: 'We want the analogy of Space ... to provide for something like the idea of absence and presence—but not just of absence and presence in the most utterly general sense these words could bear, but absence or

The upshot of the weaker point is this: if the requirement deduced in the first stage of the argument can be satisfied in Strawson's auditory universe, the second stage of the argument cannot be completed. This difficulty stems from the quite unparalleled role played by the master-sound in generating the space of the auditory world. Since change of position is not logically tied to a change in the subject's relations to the occupants in the space, but rather is a change in one particular aspect of his experience, there is no reason why that feature of experience to which it is tied should mimic the dimensionality of space.

In addition to wrecking the argument, the master-sound is the source of extreme disanalogies with our own system of spatial relations. The space of Strawson's auditory universe is an absolute space, and not a framework constituted by the spatial relations of its occupants.<sup>8</sup> The fact that 'same place' in Strawson's auditory universe is not dependent upon 'same thing' is not merely a curiosity; it lays the whole scheme open to the most straightforward phenomenalistic reduction, since 'God save the Queen is now playing at position *L*' is apparently equivalent to 'If master-sound of pitch-level *l* were heard, God save the Queen would be heard'. But, as the Carnap of the *Aufbau* learned to his cost, in a universe where space is constituted by the spatial relations between things, there can be no phenomenal characterization of being at a particular position, since things can move and change.

Taking the master-sound illustration seriously, we are prevented from completing the intended defence of the Kantian thesis, or are committed to regarding as spatial a scheme of thought so radically unlike our own as seriously to undermine the interest of the conclusion, were we able satisfactorily to defend it. Moved by these connected considerations, we must surely suspect that it is the illustration that is at fault, and not the argument which it is intended to illustrate. And further

---

presence in a sense which would allow us to speak of something being to a greater or lesser degree removed from, or separated from, the point at which we are.

<sup>8</sup> It is perhaps surprising that Strawson should invent such a space in Chapter Two, having emphasized in Chapter One the mutual identification-dependence, in our world, of things and places—a dependence which arises, on one side, from the fact that places are not intrinsically perceptible.

reflection appears to show the master-sound to be unnecessary even if an illustration of the argument is sought in a wholly auditory experience. For it may be held<sup>9</sup> that a one-dimensional space which is genuinely parallel to our own may be constructed out of an auditory experience (or, indeed, out of any experience) provided that experience exhibits such order and regularity that its course can be seen as simultaneously due to the way the world is laid out and to the subject's continuous motion through it. Provided the subject's experience is sufficiently regular to enable him to establish short-term generalizations of the form:

An experience of kind *k* will intervene between any experience of kind *k'* and any experience of kind *k''*

it might appear that he can distil from its changing course a more or less detailed map of his world, with an object of kind *k* between (in a 'travel-based' sense of this word) objects of kind *k'* and *k''*. Possessed of such a map, the subject can make empirical sense of the distinction between a change in his position and a change in the world, and being able effectively to apply this distinction, he can make revision of, and additions to, his map. That the resulting theory of the world has an interlocking, holistic character, as the subject simultaneously solves for the way the world is laid out and for the route he has taken through it, is no objection since it might reasonably be taken to mirror a feature of our own scheme. The holistic character of the resulting theory certainly renders it immune to any simple phenomenalist reduction.<sup>10</sup>

More important for our immediate purpose, this way of introducing a spatial order into an auditory world seems to provide a much better illustration of the themes of Strawson's argument. Distinct but qualitatively indistinguishable sounds can now be distinguished by their positions in the 'travel-based' ordering, and, no matter how wide a segment of experience is taken, distinct sounds can present genuinely indistinguishable appearances. Furthermore, the second stage of the argument is not undercut since we run no risk of dimensionless parallels. To

<sup>9</sup> It has been held by Jonathan Bennett, in his *Kant's Analytic* (Cambridge University Press, Cambridge, 1966), p. 37.

<sup>10</sup> An illustration of this point will be developed later; see pp. 92-4.

construct a travel-based space is necessarily to construct an *ordering* of the objects or phenomena whose relations constitute the space.

In view of these considerations, let us suppose that Strawson would accept this as a better illustration of his argument. Eliminating the master-sound enables us to concentrate upon the nub of the argument—the connection between objectivity and reidentification. Here, too, I believe there is reason for scepticism.

My first objection, which does not run very deep, concerns the role which the concept of reidentification plays in the argument. The theories with which Strawson is concerned incorporate the idea that the subject has experience of phenomena which are independent of his experience of them. From this general acknowledgement of the independence of the world, Strawson is surely entitled to extract the corollary that the temporal dimensions of an experienced phenomenon may diverge from the temporal dimensions of any experience of it. Further it seems reasonable to suppose that Hero must be able to understand one particular application of this general idea, namely the possibility that the experienced phenomena should continue throughout a gap in his experience of them. Strawson expresses this idea in the language we use to talk about the persistence of material bodies, so that Hero is to think that he has experience of auditory *items* which continue to *exist* while unobserved and which may therefore be *reidentified*. But it is not clear that the concept of identity need be involved here at all, still less that it need be involved in just the way it is involved in our scheme of three-dimensional bodies.

In the first place, it seems that this particular application of the possibility of temporal divergence between phenomenon and experience can be captured with the concept of *continuity*, as it occurs in the thought that it may continue to rain after one falls asleep, or that it may rain continuously between the time one falls asleep and the time one awakes. It would appear that the idea of its raining continuously is prior to, and independent of, the idea of a *single rainstorm*, and that it is possible to enrich what Strawson calls a 'feature-placing' language with an operator having the force of 'continuously' without disturbing its ontological simplicity—without introducing quantification



over, and reidentification of, particulars. If this is so, a theory of an objective world can be couched in a feature-placing language, and the concept of identity does not belong in Strawson's argument at all.<sup>11</sup>

Aside from this, there is a worry which arises even if we suppose that our Hero does express the idea of continuous sound by using the concept of identity. Let us suppose that Hero registers the independence of the world by allowing for the possibility of there being later, unheard parts of the same sound of which he has heard an earlier part, and therefore for the possibility of his hearing still later parts of the same sound of which, before an interruption, he has heard an earlier part. This is still not to think of an auditory *item* which *persists through time*, but rather of an auditory *process* which *is extended in time*. If the concept of reidentification is to be used in connection with processes, it must be understood that it is being used in a different sense from that which it has in connection with things. We reidentify a process when we hold that an occurrence encountered at one time is *part of* the same process as an occurrence encountered at another, but it is a distinctive (and some have thought incoherent) feature of our conceptual scheme of material bodies that we suppose an object to be both present *as a whole* on one occasion, and literally identical with an object present *as a whole* on another.

Now, a concept of reidentification can be used in connection with processes, and Strawson's detailed discussion of the auditory universe makes it clear that this is the concept which he intends.<sup>12</sup> Nevertheless, in the absence of any explanation or qualification, the use of the concept of reidentification which was originally introduced in the description of our conceptual

<sup>11</sup> In view of Strawson's defence of the coherence of a feature-placing scheme of thought later on in the book (pp. 202-13), it is perhaps surprising that he gives the impression in this chapter that he has demonstrated that any coherent scheme of thought about the objective world must involve the idea of reidentifiable particulars. Strawson does occasionally appear to restrict his question to schemes of thought involving particulars (as in: 'What are the most general statable conditions of knowledge of objective particulars?'—p. 72) but to defend the formulation of the argument by taking this restriction seriously, is to do so at the cost of significantly reducing its interest.

<sup>12</sup> '... identified as *part of* the same particular M as that of which the previously heard instance of A was a *part* ...' (p. 70); 'There is a clear criterion for distinguishing the case of hearing a *later part* of a particular auditory sound-sequence of which the *earlier part* has been heard previously ...' (p. 77) etc. etc. (My italics.)

scheme of material bodies carried with it the suggestion of a greater parallel between that scheme and the scheme of the auditory universe than is strictly warranted—a parallel which can only be purchased at the cost of ignoring the considerable difference between things and processes.

Identity is, indeed, a double irrelevance, since it appears not only that continuity can be registered without it, but also that identity can be recognized in the absence of continuity—at least where processes are concerned. We can quite intelligibly hold that later occurrences are parts of the same extended game as earlier occurrences to which they are not linked by any continuous series of game stages. This point has been raised in objection to Strawson's argument, for it seems to open the possibility of thinking reidentifyingly without thinking objectively.<sup>13</sup> The obvious reply is that the criteria of reidentification mentioned in the argument must be restricted to those which require continuity; but if this reply is made the concept of identity once again drops out of the argument, in favour of that of continuity.

The reason why I said that this objection was not very deep was that it seems simply to invite a restatement of the argument in terms of the concept of continuity. Hero must be able to make sense of the idea that, after his perception has ceased, there should be  $\phi$ -ing stages continuous with the  $\phi$ -ing stages which he perceived, and therefore that some such stage may be encountered later. But not just any later  $\phi$ -ing stage is continuous with a given earlier one. So, Hero must have a way of drawing a distinction between those later  $\phi$ -ing stages that are and those that are not, continuous with a given earlier stage. Is this distinction not exactly parallel with our distinction between qualitative and numerical identity, and will it not similarly presuppose space?

I do not believe that this argument is successful, for it seems to beg the question against a No-Space world in a subtle but decisive way.

In a spatial world there is no absolute notion of (temporal) continuity; we can only speak of spatio-temporal continuity. Now, in order to affirm upon the basis of a later-perception of

<sup>13</sup> See Don Locke, 'Strawson's Auditory Universe', *Philosophical Review* lxx (1961), pp. 18-32.

$\phi$ -ing that the  $\phi$ -ing one experienced at  $t$  did continue (did have later stages) one has to be sure, not merely that the later  $\phi$ -ing is continuous with *some*  $\phi$ -ing in existence at time  $t$ , but also that it is continuous with the particular  $\phi$ -ing experienced. For, in a spatial world, and possibly only in a spatial world, there can be distinct but simultaneous instances of the same universal. Thus, to be sure that the toy I saw at time  $t$  survived until time  $t'$  it is not enough to be sure (i) that the indistinguishable toy I see at  $t'$  was in existence at time  $t$ , one must in addition be sure (ii) that the route by which it arrived at its position at  $t'$  was one which started from its occupation at  $t$  of the position in which I saw a toy.

Compare this with a very much simpler way of thinking—one which does not admit of the possibility of distinct but simultaneous instances of the same universal. The objectivity of  $\phi$ -ing is recognized in this scheme (that is to say, it is intelligible that it be  $\phi$ -ing when no  $\phi$ -ing is perceived) but if, at any time,  $\phi$ -ing is perceived, then that is all the  $\phi$ -ing that the universe affords. Justice is done to this conception of reality by utterances of the unrestricted form 'It's now  $\phi$ -ing'. Now, if Hero thinks in these terms, he will certainly make sense of the idea that the  $\phi$ -ing that he is perceiving might continue, but by this he means no more than that it may  $\phi$  for all times between the time at which his experience of  $\phi$ -ing ceases, and some later time. And while in the spatial scheme he would have two questions to ask upon encountering  $\phi$ -ing after a gap, now he has just one: 'Was this  $\phi$ -ing going on continuously between the time I ceased, and the time I began, to experience  $\phi$ -ing?', for by this he means: 'Was it  $\phi$ -ing for all times between the time I ceased, and the time I began, to experience  $\phi$ -ing?'. In this crude way of thinking there is no parallel to the second question of the spatial scheme, no analogue to the distinction between qualitative and numerical identity, and therefore no need for a criterion employing quasi-spatial considerations to assist Hero in drawing it.

If this is correct, then the space Strawson extracted out of the concept of objectivity is the space he smuggled into it, by limiting his attention to those theories of the objective that allow for distinct but simultaneous instances of the same universal. It is not surprising if such theories can be shown to be

implicitly spatial; it is precisely for this reason that we were prepared to allow that the second stage of the argument might be completed. Perhaps there is some hidden incoherence in the crude and limited way of thinking, but that has to be shown, and if it can be shown the Kantian thesis can be established directly and Strawson's argument becomes an unnecessary circuit.

It is true that there is a distinction which someone thinking in these very crude terms must understand: namely, between the case where it is, and the case where it is not uninterruptedly  $\phi$ -ing during a gap in his  $\phi$ -experience. For in this sense, it remains true that 'not just any later  $\phi$ -ing stage is continuous with a given earlier one'. But space does not appear to be involved in this distinction in any obvious way, and if it is involved in some unobvious way, this also has to be shown. It certainly cannot be shown by gesturing towards a distinction in which space is involved—the distinction between qualitative and numerical identity—but to which the required distinction is not remotely analogous.

To defend the Kantian thesis, the idea of space must be shown to be implicitly involved in the very idea of existence unperceived, even as it is embedded in such a purported scheme as this. It is possible to find in Strawson's chapter materials for another line of defence of the Kantian thesis—a line of defence which would have just this effect.

## PART TWO

Strawson suggests that thinking of an auditory experience as experience of an objective world confronts our Hero with the problem of 'making sense of' the idea of sounds existing unperceived. He maintains that 'the most familiar and easily understood sense in which there exist sounds which I do not now hear is this: that there are places at which those sounds are audible but at which I am not now stationed.'<sup>14</sup> Space is clearly one way in which this difficulty can be resolved. Various other ways in which we 'make sense of' the idea of unheard sounds are mentioned, namely those which bring into play the idea of one sound drowning another, and the idea of deafness, but Strawson maintains that Hero cannot make use of them.

This is obviously a sketch of a line of argument rather than the argument itself, and in the text it is woven together with the argument we have just considered in a way which makes it difficult to disentangle, but I think that it is interesting and distinct, and I shall try to elaborate it. What, then, is the problem, and why should space be thought indispensable to its solution?

Hero must be able to understand the hypothesis, even if, in fact, he never believes it to be the case, that the phenomena of which he has experience should occur unperceived. Now, the idea of unperceived existence, or rather, the idea of existence now perceived, now unperceived, is not an idea that can stand on its own, stand without any surrounding theory. How is it possible that phenomena of the very same kind as those of which he has experience should occur in the absence of any experience? Such phenomena are evidently *perceptible*; why should they not be perceived? To answer this question, some rudimentary theory, or form of a theory, of perception is required. This is the indispensable surrounding for the idea of existence unperceived, and so, of existence perceived. (It is not to be thought that the idea of existence unperceived is an additional hurdle to be surmounted after the idea of existence perceived has been understood; the two ideas are sides of a single idea: the idea of an objective world.)

The same point can be put in other words. We might pretend for a moment that we are tracing the development, in a child's thought, of an utterance 'It's  $\phi$ -ing', originally tied to a recurring pattern of his experience—a cry with which experiences of a certain kind are greeted.<sup>15</sup> For an utterance like 'It's  $\phi$ -ing', originating in this way, to become an assertion about an objective world, it must loosen its tie with experience, so that it makes sense to suppose that it is true even when no experience occurs. But, although it must loosen its tie with experience, the tie must not be severed; that which is potentially true in the absence of any experience must be the very same statement as may, on occasion, be affirmed upon the basis of experience. There must be no question of allowing for 'It's  $\phi$ -ing' to be true in the absence of experience by introducing a new sufficient

<sup>15</sup> I do not mean to suggest that this is the way all concepts of the objective world originate; far from it. See Part Three below.

condition for its truth, unconnected with its existing basis. This would merely produce ambiguity, so that what is required would not yet have been accomplished—sense has not been made of the idea of the very same state of affairs that is on occasion experienced obtaining in the absence of experience. Now, we can detach 'It's  $\phi$ -ing' from experience, without pulling the concept apart, only if that in virtue of which 'It's  $\phi$ -ing' is true is connected with experience by some condition which is sometimes, but not always, satisfied. The proposition 'It's  $\phi$ -ing' will then be understood to entail that, if that condition is satisfied, it may be perceived to be true. In the formulation of the condition there lies a theory, or the form of a theory, of perception.

Provided that he is capable of telling whether or not this condition is satisfied, such a connection with experience allows Hero to give empirical content to the supposition that it is now  $\phi$ -ing, irrespective of whether he currently perceives that it is  $\phi$ -ing. If it is true that it is now  $\phi$ -ing, then it must be the case that if the condition is satisfied, he will perceive it to be  $\phi$ .<sup>16</sup>

This, then, is what it is to 'make sense of' the idea of existence unperceived. And the requirement that Hero have a conception of the world of sufficient complexity to enable him to understand why what is perceivable should sometimes be, and sometimes not be, perceived will surely rule out some purported theories of the world of excessive simplicity. But have we any reason for thinking that it will rule out all but the spatial theories; that 'the most familiar and easily understood sense in which there exist unperceived phenomena is the only sense'?

There are two kinds of explanation of why a perceptible phenomenon may not be perceived, if we exclude those that obviously rely upon spatial notions, such as the observer's being in the wrong position, or having the wrong orientation, or

<sup>16</sup> Do not say: empirical sense has not been given to the supposition that it is  $\phi$ -ing unperceived—i.e. that it is  $\phi$ -ing and the opportunity to establish whether or not it is  $\phi$ -ing is not taken up. All that can be required is that empirical content be given to the hypothesis that it is  $\phi$ -ing, and in such a way that it may be the case that it is  $\phi$ -ing unperceived. The supposition that it is  $\phi$ -ing unperceived is one whose intelligibility is consequential upon such a way of giving sense to the simple supposition that it is  $\phi$ -ing, but it is obviously not in its turn a supposition that needs to be given empirical content in the sense of conditions under which one can conclusively establish that it obtains.

there being something in the way. There are those that cite deficiencies in the perceiver, such as that he is inattentive, unreceptive in the proper modality, unconscious, or asleep. And there are those which cite the absence of factors in the world which are causally necessary for perception, as the absence of light is cited to explain why we cannot see a table. For reasons which I try to explain in the next section, explanations of this second kind do not represent a very promising avenue for exploration, if we are trying to find a non-spatial way of making sense of existence unperceived. Very briefly: we can make sense of the idea of a *material* object or substance existing in the absence of conditions causally necessary for its perception, but we cannot do the same for sensory objects; a rainbow cannot exist in the darkness, even if, were there to be light, a rainbow would be visible. And it appears that, if our Hero is to think of his experience as of a world, and this world is not to be a spatial world, it will be a world composed of phenomena analogous to our sounds, smells, and rainbows, rather than to our material substances.

But there does not appear to be any parallel reason why Hero should not make sense of the idea of unperceived sounds (or, more generally, phenomena) by thinking in terms of some block of unreceptivity in himself. 'Perhaps', he thinks, 'there are sounds which I do not now hear, because I am unreceptive; if I was to become receptive, I would be able to hear them.'

Let us suppose that this abstract form of a theory is filled out in the following way. Let us suppose that Hero's prior experience had been of the unceasing auditory sequence 'tick tick tick' but that upon one occasion, the sequence in experience had been 'tick tick      tock tock ...'. Why should Hero not use ordinary canons of scientific inference (let us not inquire too closely into what they are) to hypothesize that there was a tick which he did not hear, and understand this hypothesis in turn by supposing that he must have been unreceptive?

The objection Strawson would make to the coherence of this scheme of thought must be gathered from the following extract drawn from the passage in which he discusses various ways in which we make sense of the idea of unperceived sounds: 'Alternatively, they turn upon such an idea as that of failing sensory powers. But why do we think of our powers failing rather than the

world fading? This choice cannot be used to explain a conception it presupposes.<sup>17</sup>

I am not at all sure what objection Strawson has in mind here, but perhaps it is this. 'If one asks oneself why, in any particular case, one supposes that one's sensory apparatus is defective, it is clear that such a judgement cannot rest only upon internal features of one's experience (e.g. hearing nothing, or hearing things fainter and fainter), since it cannot be logically ruled out that there should be nothing to hear, or that what there is to hear is getting fainter and fainter. A judgement that one's sensory apparatus is defective must rest upon a view that this or that thing is there to be heard. Since the view that one's apparatus is defective ('this choice') must rest upon ('presuppose') a view as to what objectively exists, it cannot be used to give the indispensable surrounding ('explain a conception'), which the idea of what objectively exists was discovered to require.'

If this is the objection, then it appears wrong in principle. It is right to insist that all of the elements of the theory of an objective world should be present, but wrong to insist that they be independently intelligible. It is true that the idea of a perceptual breakdown presupposes the idea of an objective reality, and that, upon the envisaged scheme at least, the idea of an objective reality presupposes the idea of a perceptual breakdown (or lack of receptivity). The ideas form a circle, and any theory constructed with their aid will have a holistic character as a result. Propositions about how the world is will be derivable from propositions about the course of Hero's experience only when they are taken together with propositions about when he was, and when he was not, receptive, while propositions of this latter kind will in their turn depend both upon propositions about what Hero is (or is not) experiencing, and also upon propositions about what there is to be experienced. Hero must see the course of his experience as simultaneously determined by the way the world is and his changing receptivity to it; each is connected to experience, but only as modified by the other. All this is correct. What is not correct is that there is anything objectionable in principle in such an arrangement.

<sup>17</sup> *Individuals*, p. 74.



The best possible reason against objecting to such a structure in Hero's theory is that it can also be discerned in a spatial theory. In the modified version of a spatial scheme in the auditory universe, Hero can tell that he has changed position by the changing course of his experience, but only when this is taken together with a map of a fairly stable world. But that map, in its turn, can only have been established, and must constantly be revised, by Hero's adopting views as to where and when he is moving. (In the somewhat Cartesian setting of the auditory universe, in the absence of other subjects to whom Hero and his movements can be objects of perception, all it can mean for him to be at a position is for him to perceive what is audible at that position.) The parallel between these two kinds of theories is not complete, since, in the place of an absolute notion of receptivity—present or absent at a time—the spatial theory effectively employs a relativized notion: receptive to (= located at) this or that position. But this greater complexity does not prevent its central concepts from having that interlocking character which would appear to expose it to Strawson's criticism, if that criticism was well founded.

In fact, a stronger point might be suggested, namely that it is not merely permissible, but positively necessary, for that condition which is to account for the presence or absence of perception to be connected *a priori* with, and therefore, known to be satisfied only upon the basis of, propositions about the way the world is. For it was precisely this feature of the revised spatial theory of the auditory world which secured for it an immunity to that simple phenomenalist reduction which threatened the master-sound theory. (In the master-sound theory, the subject's changing position is definitionally tied to a change in one phenomenally identifiable aspect of his experience.)<sup>18</sup>

Although each person in a large circle of people can be sitting upon the knees of the person behind him, this is not a feat which only two or three people can manage. Perhaps the objection is not that there is a circle in a theory relying upon deafness or unreceptivity to give sense to the idea of existence unperceived, but that the circle is too small. For, while it is true in the spatial theory that deciding whether or not one has moved (and there-

<sup>18</sup> I shall discuss the significance of this irreducibility below; see pp. 124–125.

fore deciding whether a change in one's experience signals objective change at some given place) requires taking as given certain propositions about the way the world is, these are not the very propositions about the world whose truth one is required to establish, but rather propositions about how it is with adjacent places. (This reveals another part of the structure of the theory: the subject can only move continuously through space). Of course doubt might be raised about the condition of these adjacent places, which could be resolved in the same way provided weight is shifted on to knowledge of still other places; an indefinite series of such challenges could bring us back to the place from which we started. However, the theory, though interlocking, has enough structure to get off the ground; one who holds it can meet a challenge parallel to the challenge in Strawson's rhetorical question:

But why do we think our position changing rather than the world changing?

Compare this with the scheme using receptivity. Hero is supposed to be able to make sense of the idea of its now  $\phi$ -being unperceived by relying upon the concept of receptivity, and in particular, by supposing that, if he were now to become receptive, he would perceive  $\phi$ -ing. However, this only gives content to the idea that a sound now exists unperceived if there is some criterion of Hero's now becoming receptive, other than his perceiving  $\phi$ -ing. But what could it be?

Equally, Hero was supposed to be able to understand the hypothesis that there was an unperceived tick by using the supposition that he was unreceptive. But the past regularity in his experience cannot be regarded by Hero as conclusively establishing the hypothesis that there was an unperceived tick. If it does, this would not be because Hero had made the uniformity of nature a logical truth – no one could do that. It would simply be that he had established a new, and independent, sufficient condition for the statement that there is a tick, and therefore failed to give sense to one and the same state of affairs existing both perceived and unperceived. But, if the inductive considerations fail conclusively to establish the hypothesis of an unperceived tick, then Hero must be able to distinguish, at least in thought, between the case in which the

regularity was perpetuated, and the case in which it was not. Here we should like Hero to be able to appeal to the counterfactual conditional 'If I had been receptive, I would have/would-not have heard a tick'. But such a conditional is quite vacuous if the only possible conception that he can have of his being receptive at that time is simply that of being able to hear what is there to be heard.

Here, surely, are the materials for a possible line of defence of the Kantian thesis – a line of defence which rests upon the idea that only a spatial theory can satisfy the demand that the factor accounting for the presence or absence of perception of perceptible phenomena should be at once *a priori* connected with the propositions about the world, and yet subject to significant empirical control. I shall not now pursue this line any further; to do so would involve the consideration of a variety of alternative schemes<sup>19</sup> in a detail which is not perhaps commensurate with their interest. The principle of the argument should be clear, and if it is clear, we have perhaps derived such illumination of the role of space in *our* thinking as it is in the power of this or any, defence of the Kantian thesis to afford. After all, as Strawson himself emphasizes, this is the object of the exercise.

Instead, in the remaining two sections, I want to turn to what must be a brief, and I fear rather dogmatic, consideration of the question of whether a coherent theory of an objective world can be constructed upon the basis of an experience that is wholly auditory, even when that experience exhibits whatever degree of order and connectedness is necessary for the subject to apply 'travel-based' spatial notions to it. There does not appear to be any concept parallel to that of matter or material substance, which can be framed in the auditory universe. The first question I want to ask is: Can there be a world without substance?

### PART THREE

It seems possible to draw a distinction between two kinds of properties which objects may have, though a complete elaboration

<sup>19</sup> A spatial scheme is not the only scheme to employ a relativized receptivity condition, with the possibilities of additional empirical control which that provides; we can make sense, perhaps, of the idea of being *co-receptive*, where receptivity is relativized to a universal. And there are other possibilities.

tion and defence of the distinction would be a very difficult task. Into the first kind fall those properties which are dispositions to affect sensitive beings with certain experiences – these we might call *sensory properties*, or, in deference to a long-standing tradition in philosophy, *secondary properties*. For an object to have such a property is for it to be such that, if certain sensitive beings were suitably situated, they would be affected with certain experiences, though this property may, in its turn, be identified with what we should normally regard as the ground of the disposition. However, in the first instance, a sensory property is a dispositional property.

It is not necessary for our immediate purpose to have any other characterization of primary properties than as non-sensory properties of objects. So defined, the class is extremely heterogeneous. What is important, though, is that the properties constitutive of the idea of material substance as *space-occupying stuff* should be acknowledged to be primary. These include properties of bodies immediately consequential upon the idea of space-occupation – position, shape, size, motion; properties applicable to a body in virtue of the primary properties of its spatial parts; and properties definable when these properties are combined with the idea of force (e.g. mass, weight, hardness). The way these properties relate to experience is quite different from the way sensory properties relate to it. To grasp these primary properties, one must master a set of interconnected principles which make up an elementary theory – of primitive mechanics – into which these properties fit, and which alone gives them sense. One must grasp the idea of a unitary spatial framework in which both oneself and the bodies of which one has experience have a place, and through which they move continuously. One must learn of the conservation of matter in different shapes, of the identity of matter perceived from different points of view and through different modalities, and of the persistence of matter through gaps in observation. One must learn how bodies compete for the occupancy of positions in space, and of the resistance one body may afford to the motion of another. And so on.

To say that these primary properties of matter are theoretical is not to explain or to mystify, but to highlight an analogy between the way our grasp of them rests upon implicit

develop and apply these concepts independently of our having all the corresponding sensations.<sup>26</sup>

With such an excellent and detailed discussion to appeal to, I feel easier in giving nothing but a very brief sketch of the distinction.<sup>27</sup>

All it can amount to for something to be red is that it be such that, if looked at in the normal conditions, it will appear red. This formulation embodies what we might call the dispositional route from subjective experience to objective property, a route with which we are already familiar.<sup>28</sup> Philosophers have tried to provide a different account of what it is for a colour ascription to be true which does not so much involve a different route, as an attempt to make the most direct possible leap from subjective experience to objective property. They have tried to make sense of the idea of a property of redness which is both an abiding property of the object, both perceived and unperceived, and yet 'exactly as we experience redness to be'. By concentrating upon one's experience of colour, one is supposed thereby to know what it is for an object to have this property: 'This', one is to say, referring neither to the experience nor to any primary property of the thing, 'this, just as it is, can exist in the absence of any observer'.<sup>29</sup>

But the leap gets us nowhere, for it inevitably involves an attempt to make sense of an exemplification of a property of experience in the absence of any experience. Wittgenstein once imagined a world in which there were places which affected everyone painfully, so that pains were located at places in the

<sup>26</sup> Norman Daniels, *Thomas Reid's Enquiry* (Burt Franklin, New York, 1974), xiv; see also Ch. IV.

<sup>27</sup> This way of drawing the distinction also echoes recent work on the interconnection between the primary properties and the idea of space see A. M. Quinton 'Matter and Space', *Mind*, lxxiii (1964); on the idea of secondary qualities as dispositions to affect us with experiences see J. Bennett, *Locke, Berkeley, Hume* (Clarendon Press, Oxford, 1971), Ch. IV. I differ from Bennett in not making the dispositional character of the secondary qualities a matter of the meaning of sentences ascribing secondary qualities, but relying instead upon the obscurer notion of that in which their truth consists. It seems decisive against any dispositional account of the meaning of such a term as 'red' that the only way to characterize the experiences red objects produce in us is as such.

<sup>28</sup> See pp. 88-9 above.

<sup>29</sup> See, for example, J. L. Mackie, *Problems from Locke* (Clarendon Press, Oxford, 1976), Ch. 1. Mackie regards such a concept as intelligible, and used by the common man, though in fact Mackie himself does not believe there is any scientific use for it. In these views, he follows Locke.

way we locate smells. Suppose this fantasy came true. Would it then make sense to give a non-dispositional account of what it is for there to be a pain at such and such a spot; to suppose a 'pain as we feel it' existing in the absence of any observer? What can the latter form of words mean save that something awful is going on there, and how can that be, when there is no one who is hurt? To modify a dictum of Wittgenstein, conceiving of a pain which no one feels upon the model of a pain which one does feel is none too easy a thing to do.

We may ask a philosopher who claims to find intelligible the idea of an objective property extracted from our experiences of colour in this direct way, whether or not such a colour property can characterize an object in the dark. He can hardly say 'Yes', since it would be quite obscure how a 'colour-as-we-see-it' can exist when we cannot see it, and how our experiences of colour would enable us to form a conception of such a state of affairs. Further, it would have to be explained in what the difference between such an objective colour property, and the dispositional property, consists. Observing the results of switching on the light merely tests for the dispositional property; what could show whether or not objects did in fact retain these other colour properties in the dark? To maintain, on the other hand, that such colour properties cannot be true of objects in an unlit cellar seems to undermine the status of the property to being an objective property of a body, since it seems to depend for its existence upon the conditions necessary for the human perception of it. Further, the concept is said to be different from a dispositional property but it is difficult to see in what a grasp of the supposed residue would consist. Presumably, it is conceivable that objects which are not really 'red-as-we-see-them' should appear red to us; indeed, this appears to be the situation Locke supposed actually to obtain. But what one conceives, when one conceives that objects which appear red to us are, in addition, really red, or are, in addition, not really red, and how one might manifest, either verbally or behaviourally, these supposed conceptions, is quite opaque.<sup>30</sup>

These remarks are equally directed against those 'hard nosed' philosophers who wish to maintain that "science has shown that objects are not really red". Such a position would equally require the intelligibility of a non-dispositional concept directly fashioned from experience, which I am trying to deny.

suppose a disposition like fragility has in the arrangement and binding of molecules; more relevantly, we suppose that any disposition of a place to affect us with certain experiences has a ground of this kind in the occupation of that place by a matter of a certain sort. Now, Hero can have no idea of such a ground for the dispositions which places have to affect him in a certain way. He has no resources, or at any rate has not obviously been provided with resources, for forming the idea of any property of the world that is not a disposition of the world to affect him in a certain way.

It is important to appreciate the difference between these two kinds of ground for a conditional if we are to gain a proper understanding of phenomenalism. Sir Isaiah Berlin has objected to phenomenalism that it reduces categorical existence to the truth of subjunctive conditionals which are not, in their turn, grounded in anything else.<sup>33</sup> Dummett has replied, on behalf of the phenomenalist, that, provided he gives up bivalence for statements concerning material objects in remote parts of the world, there is no reason why he should suppose that any subjunctive conditionals are barely true, since he can maintain that propositions about the explored material world are true in virtue of observed regularities in our experience. But it is clear that Dummett is not offering what Berlin was missing, namely a ground of the second kind for these subjunctive conditionals—a relatively abiding property of an object or place which, together with a subject's presence, could be used to explain his experiences. Whether or not it is something that would or should worry the phenomenalist, Berlin certainly put his finger upon a deep conceptual prejudice of ours that is offended by dispositional properties without categorical grounds of the second kind. The sense of disquiet which we feel at the idea of two glasses which are exactly alike in all that is abiding, yet different in that if one is struck it will emit middle C, and if the other is struck it will not, is not at all diminished by citing the generalization upon which the proposition may be asserted, namely, that whenever in the past one has been struck, it has emitted middle C, and whenever in the past the other has been struck, it has not. And this prejudice is equally offended by the idea of two places alike in what

<sup>33</sup> I. Berlin, 'Empirical Propositions and Entailment Statements', *Mind*, *lix* (1950).

occupies them between visits, yet of which one is such that if one goes to it, one will have certain experiences, and the other is not.

This is just the situation that Hero must accept in his world; places have powers that cannot be identified with anything continuously occupying them, so that going to a place is just a basic, causally relevant factor in the explanation of the course of his experience. However, I do not want at this point to examine whether there is more to our resistance to such an idea than mere prejudice, but rather turn to this question: if this is the situation, can we continue to suppose that Hero has a coherent theory which incorporates the idea that he has experience of an objective world?

The notion of objectivity arises as a result of conceiving a situation in which a subject has experience as involving a duality: on the one hand, there is *that of which there is an experience* (part of the world) and, on the other, there is *the experience of it* (an event in the subject's biography). We have been exploring the consequences of this duality, especially the consequence that, though the temporal dimensions of these two elements overlap, they need not coincide. And, if the situation does comprise these two elements, they are not unconnected—they are not two distinct states of affairs existing simultaneously by accident or as the result of pre-established harmony. Thus unconnected, the one could not be regarded as an experience of the other, as a way of gaining knowledge of it, and thereby, of the world of which it is a part.

Now, can these features be recapitulated in Hero's scheme—in a scheme where that in virtue of which the 'objective proposition' is true can only be generalizations about the past course of Hero's experiences? The answer is surely 'No'. We do not have two states of affairs existing simultaneously, and related causally. All that exists at the position is Hero and his experiences. The only cause of Hero's having those experiences on going to that position is his going to that position. That which makes the 'objective proposition' true cannot be cited as a cause. If it is regarded as a barely true dispositional property of a position, then it is ineliminably characterized in terms which logically connect it to the event that is to be explained. It is no improvement to consider the 'objective proposition' not as



This is the first respect in which I believe that Strawson does not provide the subject of auditory experience with a coherent conception of external reality—the conception is one constructed exclusively out of sensory concepts. However, even if it is now clear that such a conception of an independent reality is not possible, it may be less clear why the conception by a subject of auditory experience of his world must take this form. After all, I have stressed that our ideas of matter are independent of any particular kind of experience, for example, tactual experience. So why can there be no analogue to the idea of matter in an auditory universe?

The issues raised by this question are enormous, and I should not be able to deal with them in this paper, even if I knew how. If the hypothetical theory is to follow ours at all closely, sounds would have to *occupy* space, and not merely be located in it, so that the notions of force and impenetrability would somehow have to have a place, and we may well wonder whether we can make sense of this without providing Hero with an impenetrable body and allowing him to be an agent in, and manipulator of, his world. But perhaps this is the wrong line to pursue. Perhaps we should explore the possibility of a theory more closely analogous to the field theory of some physicists, or even a physical theory working on principles quite different from any we have knowledge of. Unsurprisingly, I cannot myself conceive of such a theory nor do I know how one might set about demonstrating its possibility, nor whether a search for such a demonstration is even coherent. What I hope to have pointed out is the need for some physical theory over and above the ideas which Strawson attributed to his subject; not that such a supplementation cannot be provided. Though my conclusion is limited, I believe that it is worth drawing. For it is extremely tempting, upon first reading Wittgenstein's fantasy about locatable pains, to think: "How simple! This is all that separates the inner from the outer—this is all that is required to make an object, and hence a world, out of experience." Liberat-

---

something which is identical with Hero, and, since it exists unperceived, something which is characterized by primary properties.

Strawson made a half-hearted attempt to provide Hero with an audible body (pp. 84-5) but, since it does not have primary properties, it does not meet the need here indicated.

ing and appealing though this thought may be, it does not appear to be correct.

#### PART FOUR

I shall raise the second doubt I have about Strawson's auditory 'universe' rather obliquely, taking as my starting point a debate about the spatial concepts of the blind. For some centuries now, philosophers and psychologists have disagreed over whether and to what extent the spatial concepts of the blind are similar to those of the sighted. There have been those who have maintained that the blind do not, strictly speaking, have genuine spatial concepts at all; as Lotze puts it:

... the space of a blind man may not be so much what we mean by space, as an artificial system of conceptions of movement time and effort ...<sup>37</sup>

This position is also taken by Platner:

In reality, it is time that serves, for the man born blind, as space. Remoteness and proximity only mean to him the time, more or less long, and the number, more or less, of intermediaries which he needs in passing from one tactual impression to another.<sup>38</sup>

Occupying a middle position are those who agree that the distance concept of the blind is essentially a concept of the time and bodily movement necessary to traverse the distance, but who deny that in this fact there is any ground of difference from the sighted, since their spatial concepts have exactly the same feature. This is Mill's position.<sup>39</sup> And although I do not know of any text in which Poincaré specifically addressed himself to the question of the blind, he is committed by his general views on the concepts of space to occupying this position:

To localize an object simply means to represent oneself the movements that would be necessary to reach it. It is not a question of representing the movements themselves in space, but solely of representing to oneself the muscular sensations which accompany

<sup>37</sup> H. Lotze, *Metaphysic*, Vol. II (Clarendon Press, Oxford, 1887), pp. 272-3.

<sup>38</sup> E. Platner, *Philosophische Aphorismen* (1793), Vol. i, Sect. 765, p. 439. Quoted in J. S. Mill, op. cit., pp. 283-4. The most extreme and dogmatic version of this position is found in von Senden's book *Space and Sight* (Methuen, London, 1960).

these movements and which do not presuppose the existence of space.<sup>40</sup>

The other group opposed to the idea that the spatial concepts of the blind are *toto caelo* different from those of the sighted accept that the only genuine spatial concepts are those that are instanced in an array of simultaneously existing objects, and thus that are paradigmatically applicable to a simultaneously presented array, but deny that such concepts are inaccessible to the blind. According to this last group, while it is true that the blind must receive information about the spatial arrangement of the world successively, it is possible for them to organize the information they receive into a form in which genuine spatial concepts are used, or to which they may be applied. This point has been made with particular reference to the haptic perception of an object like a chair, too large to be encompassed by the hand, but presumably the point would also apply to the blind man's conception of the room or city in which he lives. Revesz puts the point about haptic perception like this:

... But even when the details have been touched, the total form is not yet given. The parts which are touched must be finally unified in a total form, in a complete impression. This synthesis presupposes a specific constructive process which we see in the visual sphere only occasionally ... Thinking and fantasy exert their effects together with intuition. The parts of a figure grasped haptically become fixated abstractly.<sup>41</sup>

It is difficult not to think of this synthesis in terms of the formation of an image, and this is the way in which Pierre Villey, a Montesquieu scholar who was himself blind, put the point:

The image which a blind man receives by touch rids itself very easily of the characteristics which constitute the modalities peculiar to tactual sensation ... The residue which it keeps, if it does not contain colouring which is absolutely foreign to tactile nerves, and if it be less rich than the contents of the visual image, may frequently not include any element which is not in the visual image, and may coincide very nearly with that.

<sup>40</sup> *The Value of Science* (Dover, New York, 1958), p. 47.

<sup>41</sup> Revesz, *The Human Hand* (Routledge and Kegan Paul, London, 1958), p. 26.

He acknowledges that his tactual perception of the chair is successive while visual perception is simultaneous, but goes on:

But if, an hour after feeling it, I search in my consciousness for the memory of the vanished chair . . . I do not reconstruct it by means of fragmentary and successive images. It appears immediately and as a whole in its essential parts . . . There is no procession, even rapid, of representations . . . I couldn't tell in what order the parts were perceived by me . . . What is the residue of this work? The limit towards which they tend, and which they appear to reach, is simply form.<sup>42</sup>

I do not mean to engage in this fascinating dispute now, but to get a little closer to my objective by extracting from it the distinction between two different kinds of spatial concepts which it highlights. On the one hand, we have what I shall call *serial* spatial concepts—concepts explained in terms of the succession or sequence of the subject's perceptions, and any muscular or kinaesthetic sensations accompanying these changes, whether they arise from the movement of the whole, or merely part, of the subject's body. I shall call these concepts 'spatial', but in view of the scepticism that is to follow, neither this term, nor the corresponding term 'travel-based' ought to be taken too seriously. For, as Poincaré said, the 'movements' can be characterized in terms which do not presuppose the existence of space.

Distinguished from these are what I shall call *simultaneous* spatial concepts, a notion which is much more difficult to make precise. Perhaps we may characterize them as relational concepts the situation for whose most direct application is one in which the elements related by them are simultaneously presented or perceived.<sup>43</sup> Relative to this characterization, the dispute about blind men is a dispute about whether someone who has no capacity to make the most direct application of a concept to the world might nevertheless possess it, and how this possession might be manifested.

However exactly the distinction between these two kinds of spatial concepts is to be drawn, that there is such a distinction

<sup>42</sup> Pierre Villey, *The World of the Blind* (Simplon, Marshall, Hamilton, Kent, London, 1922), p. 183.

<sup>43</sup> In view of the discussion of the concepts of primary properties at the beginning of the preceding part, this characterization must in no way be taken to suggest that such concepts can be extracted from just an experience in which distinct elements are simultaneously presented.

seems fairly clear. It seems fairly clear, that is to say, that there are two quite different ways in which, for example, the fact that three objects *a*, *b*, and *c* lie, in that order, upon a straight line might be established. Someone might be able to tell that the line connecting the objects was straight by means of the kind of bodily movement necessary to pass from one to another, and that *b* lay between *a* and *c* by means of the temporal relation between the experience of *a*, *b*, and *c*. On the other hand, someone who was able to see, might be able simply to see that such an arrangement existed. Equally, it seems fairly clear that we can identify, in these different ways of detecting spatial facts, the application of different kinds of spatial concepts, which have different presuppositions and which sustain different kinds of reasoning. Someone who had information given in, or stored with the use of, concepts of one kind, or in a form to which concepts of one kind would be directly applicable, would find certain problems easier, and certain problems harder, to solve than one who relied on concepts of the other kind.<sup>44</sup>

Armed with this distinction, let us return to the auditory universe, and ask what kind of spatial concepts its inhabitant has been provided with. So far, in fact, Hero has a theory of perhaps excessive simplicity, since it concerns a space of just one dimension, in which distance is only measurable upon an ordinal scale. It uses just one primitive spatial concept—'*x* is between *y* and *z*'—and as it was introduced, the concept is serial or travel-based. What it means to say that *x* is between *y* and *z* is simply that an experience of *x* will intervene between any experience of *y* which is followed by an experience of *z* and conversely.

If we were to provide Hero with analogues to our more complicated spatial concepts, such as 'arranged in a square', 'forming a circle' etc.—something that would be necessary if we were able to contemplate a generalization to a two-dimensional auditory universe—then we should have to pro-

<sup>44</sup> In fact, the distinction between serial and simultaneous spatial concepts is implicit in much of the psychological literature upon spatial perception and behaviour, especially since Tolman argued for the use of the notion of a *cognitive map* in psychological explanation. For an explicit use of the distinction see F. N. Shemyakin, 'Orientation in Space' in B. G. Ananyev *et al.* (eds.), *Psychological Science in the U.S.S.R.*, Vol. I (Office of Technical Services, Washington, 1962), pp. 186–225.

vide Hero with some way of estimating the passage of time, so that a notion of distance permitting measurement upon a ratio scale could be understood in terms of the time of normal travel. (The presence of the word 'normal' is to signal that the estimate of distance, like the estimate of position, would be subject to revision in the light of considerations from elsewhere in the theory, which would therefore retain its holistic character.) The notion of a straight line could then be defined in terms of the shortest distance between two points, but the scheme would presumably be practically unworkable unless Hero could make provisional judgements of the straightness of the path he was following upon the basis of 'bodily sensations'. (In a parallel way, our holistic scheme of re-identifiable bodies and places would be practically unworkable unless we had the ability to make provisional judgements of the identity of bodies by recognizing them.) We have not been supposing that Hero has these conceptual riches, but for present purposes it would not matter if we had, since his concepts of space would remain serial.

Now, whether or not a subject in an auditory universe could have a use for simultaneous spatial concepts is a difficult question, partly overlapping with the question about the spatial concepts of the blind which I mentioned earlier. But Strawson did not suppose that the subject in the auditory universe could, and it is this that gives me the ground of my disagreement. Strawson was quite well aware of the distinction between the different kinds of spatial concepts, and he in fact elaborates an objection to his discussion which is based upon the premiss that the subject in his auditory world would not have simultaneous spatial concepts. However exactly it is that Strawson does deal with this objection he raises to himself, it is not by denying the premiss.

The objector Strawson imagines begins by pointing out that, in visual perception, we are simultaneously presented with objects in a seen spatial array:

... these simultaneously presented elements ... are simultaneously presented as being related in another respect: viz. in a respect which leads us to characterize one as being *above* or *below* or to the *left* or to the *right* of another ...<sup>45</sup>

<sup>45</sup> *Individuals*, p. 79.

existing objects, they are not obviously concepts of relations between (independently existing) objects at all.<sup>50</sup>

Against the background of this scepticism, the immunity of the 'travel-based' theory to a simple phenomenalist reduction takes on a new complexion. Any proposition 'spatially' relating specified, preceivable 'objects' is reducible in a straightforward way to a proposition about the sequence of experiences; such irreducibility as there is comes only when Hero introduces expressions referring to 'places' whose identity conditions are tied to the whole network of propositions previously mentioned, but to no one taken individually. But it is hard to believe that an ontology appropriate to a theory of an objective world is introduced by Hero's supposed version of 'It's  $\phi$ -ing at position  $p$ ', if it is not already involved in the propositions of the form 'It's  $\phi$ -ing between where it is  $\psi$ -ing and where it is  $\chi$ -ing.'

The situation is really no different from this. There is a group of currencies each actively traded against the others in a situation of floating exchange rates. The basic propositions for describing this system will be of the form ' $\pounds 1 = \$1.75$  at the end of ... day's trading'. But we can imagine the description enriched by the introduction of the idea of the *value of the £*, something which is reckoned to be increasing, decreasing or constant by means of some averaging of its relation to all other currencies. Just as in the 'spatial' case, there is sufficient *de facto* stability in the relations between most currencies from day to day to provide the background against which it makes sense to discriminate those changes in the £- $\$$  exchange rate which are due to the pound's falling, and those that are due to the dollar's rising. Now, a proposition to the effect that the value of the pound has declined is not reducible to any one proposition of the form 'At the end of day  $d$ ,  $\pounds 1 = n$  units of X currency, and at the end of day  $d + 1$ ,  $\pounds 1 = n - k$  units of X currency.' Such a proposition is not necessary, since X may be a currency moving down with the £, and it is not sufficient, since X may be moving up rather than the £ moving down.

If this does provide a parallel for the relation between the

<sup>50</sup> If the line of reasoning expressed in these paragraphs is correct, those who deny simultaneous spatial concepts to the mind are committed to denying that they have a conception of an independently existing reality at all, which is surely very difficult to accept.

basic propositions of Hero's travel-based theory, and those which mention or quantify over 'positions', it is hard to take the theory's immunity to a simple phenomenalist reduction seriously. While a genuine theory of an independent reality will be thus irreducible, not every theory thus irreducible is a genuine theory of an independent reality.

It is a little surprising that Strawson does not treat the objection to his auditory universe which we have been considering with more understanding, for the point upon which it rests is one which Strawson himself emphasized in defence of the Kantian thesis:

'... we must have a dimension other than the temporal in which to house the at present unheard sensory particulars if we are to give a satisfactory sense to their existing now unperceived ...'<sup>51</sup>

... we want an analogy of distance—of nearer to and farther away from—for only, at least, under this condition would we have anything like the idea of a dimension other than the temporal in which unperceived particulars could be thought of as simultaneously existing in some kind of systematic relation to each other and to perceived particulars.<sup>52</sup>

The objector is simply taking this point, and insisting that, if space is to provide this system of relations, it must be a space constituted by simultaneous spatial relations; that if Hero is to think of unperceived particulars existing simultaneously with, and in relation to, perceived particulars, he must have simultaneous spatial concepts, and not those that 'turn essentially upon change'.

I may have given the impression that I disagree with all the most important points which Strawson makes in his second chapter. But this does not seem to me to be so. As important as any point I have so far discussed is something implicit in the entire procedure of discussion, something implicit in what, if he would not shrink from such a word, might be called Strawson's methodology. This is the idea that the connections between the fundamental concepts of our conceptual scheme are central objects of philosophical investigation, and that exploratory pressure may have to be put upon these connections by

<sup>51</sup> *Individuals*, p. 74.

<sup>52</sup> *Individuals*, p. 75.



imagining situations radically unlike our own. (Hero must not take on a life of his own, so that speculations about him are misinterpreted as speculations 'about what would really happen in certain remote contingencies'; Hero and his 'world' are devices for 'testing and strengthening our own reflective understanding of our own conceptual structure'.) As a model of how to pursue this essentially imaginative exploration, Strawson's chapter is unsurpassed.<sup>53</sup>

<sup>53</sup> I am grateful to Crispin Wright, David Pears, and John McDowell for reading an earlier draft of this paper and offering helpful comments.

# Meaning, Communication, and Knowledge<sup>1</sup>

John McDowell

1. According to theorists of communication-intention, we should explain what meaning is in terms of the audience-directed intentions of speakers. According to theorists of formal semantics, the functioning of language can be illuminated by considering a certain sort of formal theory of a particular language: namely, a theory competent to specify truth-conditions for all of the language's indicative sentences. It seems likely that there is something to be gained from both these ways of thinking, but it is rare for much attention to be devoted to the question how they are related. However, a welcome exception is P. F. Strawson's provocative inaugural lecture,<sup>2</sup> from which the above characterizations are drawn.

Strawson's major thesis is this. The notion of truth-conditions is the fundamental notion of formal semantics, and the idea that meanings can be specified in terms of truth-conditions is all very well as far as it goes. But it does not go far enough. Further explanation of the notion of truth-conditions is called for; and this turns out to require an appeal to the characteristic conceptions of the communication-intention theorist. So while neither tradition is incorrect, the theorist of communication-intention can claim to be closer to the philosophical foundations.

In more detail, Strawson's central argument is as follows. The theorist of communication-intention can concede that 'in almost all the things we should count as sentences there is a substantial central core of meaning which is explicable either in terms of truth-conditions or in terms of some related notion

<sup>1</sup> I should like to express my thanks for help with this paper to Brian Loar, Stephen Schiffer, and especially Gareth Evans.

<sup>2</sup> *Meaning and Truth* (OUP, Oxford, 1970), reprinted in *Logico-Linguistic Papers* (Methuen, London, 1971), p. 170. Page references are to *Logico-Linguistic Papers*.

quite simply derivable from that of a truth-condition' (p. 178); and hence that theories of meaning for particular languages can hinge on the notion of truth. With its generality over languages, this concession constitutes, we might suppose, a potentially illuminating thesis about meaning in general; and it is one in which the explanatory weight rests on the notion of truth. But 'we still cannot be satisfied that we have an adequate general account of the notion of meaning unless we are satisfied that we have an adequate general understanding of the notion of truth' (p. 180). Now when we look for accounts of truth in general—as opposed to the sort of account of truth in this or that particular language which semantic theories themselves provide—the best we can come up with seems to be such platitudes as this: 'one who makes a statement or assertion makes a true statement or assertion if and only if things are as, in making that statement, he states them to be' (p. 180). If we combine this platitude with the concession that meanings can be specified in terms of truth-conditions, we arrive at the following conclusion: to specify the meaning of an indicative sentence is to specify how things are stated to be by someone who makes a statement by uttering it. (Presumably other sorts of sentence can have their meanings explained in some derivative way.) So we have arrived at the notion of the content of such centrally important speech acts as statement-making. Strawson goes on:

And here the theorist of communication-intention sees his chance. There is no hope, he says, of elucidating the notion of the content of such speech acts without paying some attention to the notions of those speech acts themselves ... And we cannot, the theorist maintains, elucidate the notion of stating or asserting except in terms of audience-directed intention. For the fundamental case of stating or asserting, in terms of which all variants must be understood, is that of uttering a sentence with a certain intention ... which can be incompletely described as that of letting an audience know, or getting it to think, that the speaker has a certain belief ... (p. 181).<sup>3</sup>

To determine the meaning of an indicative sentence is to

<sup>3</sup> This incompletely described intention is (at least in the version formulated in terms of getting the audience to think ...) a component of H. P. Grice's analysis of utterer's meaning: see 'Meaning', *Philosophical Review*, lxvii (1957), 377, and the modifications proposed in 'Utterer's Meaning and Intentions', *Philosophical Review*, lxxviii (1969).

determine what statement can be made by uttering it; and that is just to determine what belief the sentence can be used to get audiences to think the speaker has. Strawson concludes that, the concession, 'so far from being an alternative to a communication theory of meaning, leads us straight in to such a theory of meaning' (p. 182).

According to this argument, then, theories of meaning for particular languages might be essentially as theorists of formal semantics propose. But the work of explaining the concept of meaning is done, not by the central notion of such theories, the notion of truth, but by the conceptual apparatus required for the further elucidation of the notion of truth; and that is the favoured conceptual apparatus of the communication-intention theorist.

2. In Strawson's debate, the theorist of formal semantics is depicted as resisting pressure to appeal to the communicative nature of linguistic behaviour in order to underpin his interest in truth. But this misrepresents at least some of those who think about language within the tradition Strawson is considering. There is an attractive way to defend the philosophical interest of formal semantics, as applied to natural languages, which involves no such resistance.

Someone who understands a language can hear utterances in it, not just as productions of sound, but as significant speech acts. What he has is an information-processing capacity. His senses furnish him with information to the effect that people are uttering such-and-such sounds—information which is available equally to someone who does not understand the language. What is special about someone who does understand the language is that his sensory intake yields him, in addition, knowledge as to what speech acts, with what content, are being performed. The property which distinguishes him, then, would be captured by a theory with the following powers: given a suitable non-interpreting description of any possible utterance in the language—a formulation of information available equally, on hearing the utterance, to someone who understands the language and to someone who does not—the theory would enable anyone who knew it to derive that interpreting description under which someone who understands the language

would be capable of recognizing the action performed.<sup>4</sup> Such a theory would compendiously describe the extra contribution, over and above the sharable sensory intake, which his competence with the language makes to his cognitive position on any of the relevant occasions. For any possible utterance in the language, it would yield a route from a non-interpreting description of it to an interpreting description. Thus it would reveal the relations between sound and—to speak intuitively—significance which, in a sense, constitute the language. It is hard to see what could have a better claim to count as a theory of meaning for a language.

If, for the moment, we ignore moods other than the indicative, we can restrict the scope of a theory of the envisaged sort to sayings. The theory would need, then, to make someone who knew it capable of specifying, for any indicative sentence in the language it dealt with, the content of the saying which an utterance of the sentence would be taken to be by someone who understood the language. This content-specification might be made possible by a theory which met the following description: for each object-language sentence, it entails a theorem whose form we can represent schematically as 's...p', where 's' is replaced by a suitable designation of the object-language sentence and 'p' by a sentence, in the language in which the theory is stated, suitable for expressing what can be said by uttering 's'. (We shall consider what goes in place of the dots shortly.) We want to be able to see the content expressible by uttering a sentence as the upshot of contributions from repeatable parts or aspects of it. It seems inconceivable that a theory could yield specifications of content for all possible indicative sentences in a language whose complexity approached that of a natural language, unless the theory were organized in such a way as to meet something like that desideratum: the theorems would need to be deducible on the basis of the structure of the object-language sentences, in such a way that the premiss which

<sup>4</sup> There is no implication that the theory is known by a possessor of the capacity it describes. Nor need the language in which the interpreting descriptions are given be understood by him: a monoglot Frenchman can be taken by us to recognize another Frenchman as, for instance, saying that it is likely to snow. For the conception of a theory of a language expressed here, cf., e.g., Donald Davidson, 'Radical Interpretation', *Dialectica*, xxvii (1973), 313; and my 'On the Sense and Reference of a Proper Name', *Mind*, lxxxvi (1977), 159.

registers the contribution of, say, a word to a given sentence figures also in derivations which reveal its contribution to other sentences in which it occurs.

Now suppose a theory for a given language can meet these requirements by taking this shape: what the replacements for 'p' do, in the theorems, is to state necessary and sufficient conditions for the application of some predicate to the object-language sentences. One of the requirements is that replacements for 'p' are to express what can be said by uttering the relevant object-language sentences. By way of something like Strawson's platitude (§1), this guarantees the weaker claim that the extension of the envisaged predicate is that of a truth-predicate. The envisaged theory would not be false if the predicate were written 'true'.<sup>5</sup>

Reintroduction of non-indicative moods need not radically alter the picture. The task of generating sentences suitable for content-specifying can still be assigned to a component of the total theory which deals only with indicative sentences. If the object language has more than one mood, a theory competent to impose interpreting descriptions on all possible utterances in it will need to be able to classify utterances as performances of speech acts of this or that kind (assertion, question, command, or whatever). We can require the principles which effect this classification to be written in such a way that, in the case of a non-assertoric utterance, besides enabling us to identify the kind of speech act performed, they also equip us with an indicative sentence, related to the sentence uttered in such a way that the right-hand side of its theorem in the simple theory considered above—a sentence apt for expressing the content of sayings effected by uttering the indicative sentence—will equally serve (perhaps with minor syntactic modification) to express the content of the non-assertoric speech act performed by uttering a non-indicative counterpart. The relevant relations between sentences are quite systematic, and there is no reason to suppose they could not be codified in a theory.

The argument I have outlined does not purport to show that it is compulsory to formulate the central component of a theory

<sup>5</sup> Cf. Davidson, 'Truth and Meaning', *Synthese*, xvii (1967), 304, especially at pp. 310–11. Indexicality introduces a complication; but one which is surmountable, as Strawson notes (pp. 179–80).

of a language—that component which yields sentences suitable to specify the contents of utterances—as a characterization of a predicate. Nor does it guarantee that it is possible to do so; we know it is possible for languages with certain syntactic structures, but it is an open question whether we can wrench all of a natural language into syntactic forms which are amenable. But the policy of trying to do it this way, if possible, offers advantages. For it suggests a model of the kind of thing to aim at, in the truth characterizations which Tarski showed how to construct for certain sorts of formalized language.<sup>6</sup> And Tarskian truth characterizations have special attractions for those concerned with how languages relate to reality: they allow us to lay bare, in derivations of truth-conditions from semantic properties of sentence-constituents, how a sentence's bearing on the world is dependent on specific word-world relations involving its parts, and thus, to meet a version of the desideratum mentioned three paragraphs back.

3. The appearance of the debate should now be somewhat altered. A theorist of formal semantics who defends his approach on the lines I have just sketched differs from his counterpart, in the dispute which Strawson stages, in this crucial respect: he evinces no reluctance to appeal to the notion of a communicative performance in what he offers in the way of general remarks about meaning—namely, a description of the shape which a theory of meaning for any particular language might take. On the contrary, the notion of the content of a saying is centrally important in his position. If one counts as a communication-intention theorist by virtue of thinking it unavoidable, in any adequate general account of meaning, to mention kinds of action which are standardly intentional and directed towards audiences, then Strawson's battle seems after all not to be joined. But this is not to say that the theorist of formal semantics I envisage concedes, in advance, the specific outcome of Strawson's debate.

Strawson's predominant picture involves the progressive extension of an analysis. At the first stage, the theorist of formal semantics offers an analysis of the notion of sentence-meaning

<sup>6</sup> Alfred Tarski, 'The Concept of Truth in Formalized Languages', in *Logic, Semantics, Metamathematics* (Clarendon Press, Oxford, 1956), p. 152.

in terms of the notion of truth-conditions. On behalf of the theorist of communication-intention, Strawson insists that the analysis stands in need of further analysis, and that the further analysis takes us to a second stage, at which the notion of sentence-meaning is analysed in terms of the notion of the content of an assertion. Given the aim of analysis, it looks suspect to stop at this point, with an analysis in terms of an overtly linguistic notion; an analysis should decompose a notion into conceptually prior components, and the notion of assertion seems to be on a level with the notion of sentence-meaning—too much so for analytic progress to have been achieved. As Strawson sets things out, such considerations may seem to render obligatory a third stage, in which the notion of assertion is itself analysed in non-linguistic terms—as in the account of assertoric communication, in terms of intention and belief, which Strawson recommends.<sup>7</sup>

Strawson's theorist of formal semantics sticks at the first stage of this progression. Anyone who had arrived at the second stage would be ill-placed to baulk at the third; and my theorist, as noted above, does not hesitate to appeal to the notion of the content of a saying. But it would be a mistake to suppose that this places him at the second stage of Strawson's debate. For he is not, like Strawson's disputant, driven to the notion of saying in order to underpin employment of the concept of truth in a first-stage analysis of the concept of sentence-meaning. Indeed, any such suggestion reverses the order of his reflections, in which the notion of truth enters only at the end. The predicate characterized, in its application to sentences of a given language, by a theory of the sort which he envisages would be a truth-predicate; but this thesis appears, not as a purported analysis—even an interim analysis—of the notion of sentence-meaning, but as a subsequently-noticed consequence of what gives such theories their claim to count as components of theories of meaning.

It might be thought that my theorist is still not entitled to

<sup>7</sup> The full analysis would include more than the partial specification of intention quoted above. Not only would more need to be said about the intention (so far only 'incompletely described'), the idea is that a fuller account would also introduce the notion of a conventional way of executing the intentions in question (see pp. 173–5). I am sceptical about the idea that linguistic behaviour is conventional, but in this paper I shall not consider that element in Strawson's recommended analysis.



appeal to an unanalysed notion of saying, even if his doing so does not correspond to the second stage of Strawson's debate. But what seemed, in that context, to make it compulsory to attempt a reductive analysis of assertion—one which makes no use of notions essentially connected with language—was a thought about conceptual levels which is not so much as relevant if analysis is not the aim. Strawson himself remarks (though he relegates the remark to a footnote) that when one sets out to give a philosophical account of something, conceptual analysis is not the only option (p. 172). Now nothing in the position sketched in §2 commits my theorist to analytic aspirations. And we lack an argument that meaning constitutes the sort of philosophical problem which requires analysis for its solution. The ability to understand a language is an ability to know what people are doing, in the way of performing significant speech acts, when they speak in it; if there is a problem about this, it is not unfamiliarity, or resistance to comprehension, on the part of the concepts employed in saying, occasion by occasion, what someone who understands a speech act knows, but rather an initial unclarity about how the general ability can be specifically described with a degree of systematic articulation sufficient to match the systematic way in which it evidently functions. What we need, on this view, is not conceptual analysis, but a perspicuous mapping of interrelations between concepts which, so far as this exercise goes, can be taken to be already perfectly well understood.

It is a striking fact that in the mapping offered by my theorist, the concept of meaning as such does not even appear. So far from analysing the notion of meaning, he suggests the radical thought that in describing the understanding of a language we can get along without it.

4. There is, however, a different line of argument available to Strawson: one which does not depend on the idea that what we need is analysis of the concept of meaning. It seems fair to claim that a satisfying general account of how language functions should make clear the peculiar importance which language has in human lives; and it seems obvious that one could not achieve that goal without making it clear that language is essentially communicative—that speaking and

understanding are primarily the issuing and the reception of communication. Now the position outlined in §2 does not involve denying the essentially communicative nature of language; it is not liable to the devastating objections which Strawson deploys against any such denial. But there would be justice in the complaint that more needs saying about what exactly communication is. (We might hope that this would have some bearing on the question when the application of an interpreting description to a speech act is acceptable—a question to which the position of §2 requires us to have some answer.) Strawson equates assertoric communication with the fulfilment of an intention of the kind which figures in his favoured account of saying or asserting: that is, an intention which can be partly specified as that of getting an audience to believe that the communicator has a certain belief.<sup>8</sup> Here, then, we have a starting-point for an argument, quite distinct from one which appeals to the conditions for a satisfactory analysis, for accepting not just the innocuous concession to communication-intention-theories implicit in §2 and remarked on in the first paragraph of §3, but something more like the specific communication-intention theory which Strawson recommends.

But the proposed account of assertoric communication, from which this line of argument starts, is questionable.

5. According to the proposal, a partial account of what it is for a communicator (C) to communicate to an audience (A) that  $p$  by  $\phi$ -ing is as follows: C fulfils an intention that A, through awareness of C's  $\phi$ -ing, should come to believe that C believes that  $p$ .<sup>9</sup>

The form 'communicate that  $p$ ' is perhaps mildly barbarous. It may help to have in mind some plausible substitute for 'communicate'; perhaps 'get it across' will do. But if it will, the suggestion is in doubt, for this reason: it seems that one cannot get it across that  $p$  if it is not the case that  $p$ . Thus the analysis

<sup>8</sup> Strawson's formulation (p. 181) actually says 'letting an audience know, or getting it to think, that the speaker has a certain belief'. The former phrase introduces a quite different notion. See §5, and n. 12, below.

<sup>9</sup> Only a partial account. The main omission here, in comparison with the account which Strawson outlines, is a requirement of overtness in the intention. On this, see §6.

dum requires the truth of what replaces 'p', whereas the partial analysans obviously does not; and it does not seem that a more precise formulation of the analysans could repair the deficiency, so long as the operative notion remains the notion of inducing beliefs. (Simply adding 'p' as a conjunct seems too *ad hoc* to be satisfactory.)

In the mildly barbarous idiom I am considering, then, 'communicate' belongs to a class of verbs V whose defining property is that a sentence of the form 'S V's that p' entails the truth of its embedded sentence. (Compare 'disclose', 'reveal', 'convey'.) Now Peter Unger has put forward the following attractive hypothesis: what makes a verb belong to this class is that it yields us a decision as to the presence or absence of knowledge.<sup>10</sup> (Unger in fact restricts his hypothesis to lexically unstructured verbs, in order to avoid obvious counter-examples like 'guess correctly'; but it seems reasonable to suppose that the structure of 'get it across' is not of a disqualifying sort.) In the case of 'communicate', understood as tantamount to 'get it across', it is presumably the presence rather than the absence of knowledge which is determined by the true applicability of the verb. This suggests the suspicion: perhaps we should think of communication as the instilling not of beliefs but of knowledge.<sup>11</sup>

The suspicion is one which we might have wanted to entertain in any case, even without the detour through the dubious idiom. We can find it plausible, with Strawson, that the primary intention of a statement-maker, in 'the fundamental case of stating or asserting' (p. 181), is to communicate something; to get something across. But we surely ought not to find it plausible that his primary intention is to induce someone else to form a belief about his belief. Consider what seems to be the corresponding thesis about questions: namely, that the intention of an inquirer, in the most fundamental case of inquiring, is to induce an audience to induce in the inquirer a belief about

<sup>10</sup> 'Propositional Verbs and Knowledge', *Journal of Philosophy*, lxx (1972), 301; the quotation is from p. 307.

<sup>11</sup> 'Suggests' only: Unger's hypothesis would require no more than that a communicator should have knowledge (on the assumption that it is the presence rather than the absence of knowledge which is at issue). But given that the concept of communication requires that a communicator should have knowledge, it seems irresistible to suppose that what he is doing when he communicates is sharing the knowledge.

the audience's belief. Surely any such claim would be absurd. The primary point of asking questions is not to acquire beliefs about one's interlocutor's beliefs, but to find out how things are. Correspondingly, the primary point of making assertions is not to instil into others beliefs about one's own beliefs, but to inform others—to let them know—about the subject matter of one's assertions (which need not be, though of course it may be, the asserter's beliefs).<sup>12</sup>

Strawson's proposal represents a communicator as engaged in the manipulation of his audience's beliefs.<sup>13</sup> Whether the manipulation is in the audience's interest is simply left open. Of course a belief about another person's belief is sometimes a good thing to have; it might afford—if true—insight into the other person, and—if his belief is true as well—an argument to a truth about the world. But this view of communication yields no general presumption that communication, as such, is beneficial to its recipient. That depends on the goodwill, reliability, and so forth of the communicator. If communication is conceived, by contrast, as the sharing of knowledge, it will not seem an accidental fact about communication that is potentially helpful to its recipient; communication, of its very nature, confers potential benefits, whose usefulness is grounded in that interest in how things are which no agent can lack.

6. Of course an assertoric linguistic repertoire can be exploited manipulatively—to mislead rather than to inform. Moreover, a statement-maker can be honestly mistaken even if his intention is informative. This, together with the existence of speech acts which are not assertoric at all, forces us to complicate the picture somewhat, if we are to combine the thesis that communication is the sharing of knowledge with the thesis that linguistic behaviour is essentially communicative in character.

A version of the latter thesis can still be defended on the

<sup>12</sup> On the importance of the notion of letting someone know something, in an account of communication, see Strawson's article 'Intention and Convention in Speech Acts', *Philosophical Review* lxxiii (1964), 439, reprinted in *Logico Linguistic Papers*, p. 149, at p. 156 of the reprinted version. But here, as in 'Meaning and Truth' (see n. 8 above), Strawson slides (p. 157) into a formulation in terms of getting something to think something (Even if we ignore 'or getting it to think', the account in 'Meaning and Truth' is still unsatisfactory from the present standpoint, in that it restricts the content of the knowledge instilled to the communicator's beliefs.)

<sup>13</sup> I owe this thought to Thomas Ricketts

following lines. We can say that a perceptual capacity is essentially a capacity to acquire knowledge, without committing ourselves to the false claim that knowledge is acquired in every exercise of a perceptual capacity; the concept of misperception is, precisely, the concept of a defective exercise of a capacity whose non-defective exercises issue in knowledge. In a similar way, we can say that the essential character of the assertoric use of language lies in its availability for communicating, in the sense of transmitting knowledge about the subject matter of assertions, without denying the possibility of using assertions to deceive. Communication, in the sense of the transmission of knowledge about the topic of discourse, need not actually take place in every exercise of the repertoire; nevertheless it can be essential to the repertoire that it is apt for communicating in that sense. (Indeed, we might say, it is precisely by purporting to communicate in that sense that a deceiver deceives.)

If we left it at that, however, we would be simply denying—highly implausibly—that communication takes place in deceptive or misguided uses of assertoric language. So far, moreover, we have no inkling how the concept of knowledge-sharing might be extended to non-assertoric uses of language. But the implication of our minimal concession to communication-intention theories (§§2, 3) was that, simply by virtue of being intentional and directed towards audiences, speech acts were quite generally communicative. So faithfulness to the concession requires us to find a further application for the concept of communication, one in which it is not restricted to actually informative assertions but true of speech acts in general. We shall see that we can preserve the idea that communication is the instilling of knowledge.

I shall approach the topic by way of the special overtiness which is characteristic of linguistic communication.

A concept at least closely akin to the concept of knowledge-transmission applies to modes of behaviour which we can ascribe to creatures to which we would not think of ascribing intentional action. Absence of intentional action is no bar to possession of sensory capacities; that is, capacities for the acquisition of states of informedness about the environment, in the intuitively satisfactory sense of states, resulting from a

systematic sensitivity to features of the environment, which enable behaviour to be suited, in the light of needs or goals, to the way the environment actually is. Such capacities obviously have survival value. Similar survival value would attach also to any behavioural disposition whose effect was to spread, among several individuals, the beneficial results of one individual's exercises of its perceptual capacities. A bird, say, might instinctively emit a characteristic sort of squawk on seeing a predator; other birds might acquire, on hearing such a squawk, a propensity towards behaviour appropriate to the proximity of a predator (flight, increased caution in feeding, or whatever). This propensity might match a propensity they would have acquired if they had seen the predator themselves. In such a case it would be natural to regard the squawk as a further mode of sensitivity to the presence of predators, over and above more direct kinds of perception. The upshot of this further mode of sensitivity is no less appropriately thought of as possession of information (or misinformation if things have gone wrong) than is the state which standardly results from perceiving a predator. We might jibe at the word 'knowledge', but there is no risk of over-psychologizing our account of the birds—crediting them with an inner life—if we regard such behaviour as effecting the transmission of information, and hence as constituting a kind of communication.

Now a conspicuous difference between linguistic behaviour and this kind of information-transmission lies in the intentions which are overt in speech. In successful linguistic exchange speaker and hearer are mutually aware of the speaker's intentions, in a way which could have no counterpart in merely instinctive responses to stimuli.

Strawson indeed employs the notion of overtness, when he formulates a component of his favoured account of communicative intentions which I have so far passed over. In Strawson's sketch, the communicator's intention that the audience should form a certain belief about the communicator's belief is 'wholly overt', in this sense: the communicator intends this very intention to be recognized by the audience (pp. 172-3).

It is open to question, however, whether this formulation successfully captures the special transparency of intention which is characteristic of linguistic communication. Such

mutual awareness of intention as is involved in the overtness which Strawson describes could be fully achieved even though the intention which is alleged to be the communicative intention is not fulfilled: the audience might recognize the communicator's intention that he, the audience, should adopt a certain belief about the communicator's belief, but fail, or refuse, to oblige. Thus making the communicative intention overt—securing mutual awareness of it between speaker and hearer—is represented as independent of communicative success. What seems plausible, however, is this: the appropriate mutual awareness is actually what is aimed at by the speaker's primary communicative intention, so that securing the mutual awareness is not, as in Strawson's picture, a fallible means to communicative success, but rather constitutes it. This suggests the following position. The primary communicative intention is the intention, for instance, to say such-and-such to the audience. The appropriate mutual awareness is awareness that the speaker has indeed said such-and-such to the audience. Speech acts are publications of intentions: the primary aim of a speech act is to produce an object—the speech act itself—which is perceptible publicly, and in particular to the audience, embodying an intention whose content is precisely a recognizable performance of that very speech act. Recognition by an audience that such an intention has been made public in this way leaves nothing further needing to happen for the intention to be fulfilled.<sup>14</sup>

The notion of an intentional performance is more fundamental in this context than the notion of the intention to perform it. One can sometimes divine an intention to say such-and-such behind a bungled performance, but correctly executed speech acts carry their intentions on their surface; normal understanding of correct speech is not a matter of divination. And an account of the understanding of language must start with the understanding of correct speech.

Now intentions of this sort—intentions whose content is the publication of themselves—are communicative in a sense which involves no departure from the idea that communication is the instilling of knowledge. The intention to make such an

<sup>14</sup> For the idea of speech acts as intentional acts such that the recognition of the intention is its fulfilment, see John R. Searle, *Speech Acts* (CLUP, 1969), p. 47.

intention public is the intention to let an audience know what speech act is being performed. Here, then, we have our required further application for the concept of communication. The concept applies, we can now say, at two levels. At the first level, communication takes place, as before, only when information is actually transmitted about the topic of discourse. But at the second level, the information whose sharing is relevant to the question whether communication is taking place concerns, not the topic of discourse, but the nature of the speaker's intentions; and when a properly executed speech act is understood, such information is always transmitted—not only in informative assertions but also in assertions by which information is not transmitted and in speech acts which are not assertoric at all.

If we take it that the content of the intention made public in a speech act—an intention which is communicative at the second level if not at the first—essentially involves the concept of the kind of speech act in question (it is the intention, for instance, to say such-and-such), then we cannot hope for a reductive account of kinds of speech act in terms of the intentions of their performers. But once we have turned our backs on analysis of the concept of meaning (§3), we have no obvious reason to regret this renunciation.

7. Strawson's debate has the centrality of assertoric speech acts as an undisputed background; the idea is that, if we can acceptably explain the notion of the content of these central speech acts, there should be no special difficulty about the content of other sorts of speech act. Now at the beginning of §6 I considered the thesis that availability for first-level communication—transmission of information about the topic of discourse—is of the essence of an assertoric repertoire. I suggested that this thesis cannot be quickly ruled out on the score that assertions need not transmit information; but we lack, as yet, a positive reason for accepting the thesis. Perhaps finding one would be finding a way of domesticating, within a position in which reductive ambitions are renounced, Strawson's thought that 'there is no hope ... of elucidating the notion of the content of such speech acts without paying some attention to the notions of those speech acts themselves' (p. 181).

To specify the meaning of an indicative sentence, according



occurs in something like the sense in which it is the function of the heart, say, to circulate the blood. When what gets transmitted is misinformation, there has been a malfunction of a natural process. A malfunction is as such a defect (even if, on occasion, misinformation is better for its recipient than information would have been). In this case, then, the non-indifference of the distinction between truth and falsehood is simply the non-indifference of the distinction between proper functioning and malfunction. Aims pursued in communicating do not enter the story. There are no such aims, since the behaviour is instinctive. But in an account of the (no doubt rudimentary) notion of content which seems undeniably applicable in this case, the natural function of the behavioural repertoire can serve, as it were; instead; it can occupy a position analogous to the position which was supposed to be occupied, in an account of the notion of the content of an assertion, by the alleged fact that in making assertions we aim at truth.

When the communicative process functions properly, sensory confrontation with a piece of communicative behaviour has the same impact on the cognitive state of a perceiver as sensory confrontation with the state of affairs which the behaviour, as we may say, represents; elements of the communicative repertoire serve as epistemic surrogates for represented states of affairs. It is hard to see what perplexity there could be, about the application of the notion of content to pieces of behaviour of this kind, which would persist entirely unalleviated by this thought: elements of such a repertoire represent states of affairs by virtue of standing in for them in a creature's cognitive dealings with the world.

It is plausible that the assertoric core of linguistic behaviour is a descendant, now under intentional control, of the sort of instinctive communicative repertoire we have been considering. This would account for the plausibility of the suggestion that availability for transmission of knowledge is an essential characteristic of the assertoric component of a language (§6). Of course assertions can serve purposes other than the purpose of supplying information; once the behavioural repertoire is responsive to a variety of non-informative intentions, it no longer seems appropriate to think of the transmission of information as the natural function of its exercises. However, it

remains a striking fact about assertions that knowledge can be acquired at second hand. If someone knows that *p* and says that *p*, then typically someone who hears and understands him is in a position to know that *p*. It seems unpromising to suppose that knowledge by hearsay owes its status as knowledge, quite generally, to the knower's possessing a cogent argument to the truth of what he knows from the supposed reliability of the speaker. A more attractive line of thought is that the linguistic repertoire retains, through the alteration of nature involved in the onset of self-consciousness, a form of the characteristic which was essential to its pre-linguistic ancestor: in suitable circumstances (to be spelled out in any fuller elaboration of this idea) its exercises are cognitive stand-ins for the states of affairs which they represent. An assertion will actually have that epistemological role only if the circumstances are right. But all standard assertions—excluding, that is, special cases like irony—purport to have it. Thus their possession of content—their capacity for representing states of affairs—is intelligible in terms of a suitable modification of the simple idea which seemed appropriate in the case of instinctive communication.

8. I have been ignoring the belief-expressing aspect of assertion on which Strawson concentrates. I have no wish to deny the central importance, in any account of what gives language its special place in our lives, of the way it enables us to reveal our states of mind to one another. Indeed, the belief-expressing aspect of assertion would be crucial in empirically assessing a theory of a language of the sort envisaged in §2. A theory of that sort is acceptable if the interpreting descriptions of linguistic behaviour which it yields make the behaviour intelligible; the test of intelligibility is the question whether we can make sense of the behaviour, as described, in terms of propositional attitudes which the speakers' behaviour and circumstances permit us to attribute to them, and we typically make sense of assertions as expressive of beliefs. But to concede all this is not necessarily to conclude, with Strawson, that we must understand the possession of significance by indicative sentences in terms of their admitted availability for the expression of beliefs. In fact §7 suggests the possibility of a position in which things are the other way round.

properly understood. But what is needed for full exposure to the significance of the sentence is not thoughts about the mental state of the original informant, but acceptance of the further sentences which form the required theoretical context, and a capacity to employ them appropriately in connected conversation.<sup>20</sup> Once the child has worked his way into a fair competence with an appreciable amount of a language, he will surely be engaged in, and sensitive to, assertion as expressive of belief. But it is not because belief-expression is by then in his picture of his dealings with language that content, gradually more fully understood, has for some time been an appropriate constituent of our picture of his dealings with language. Content is in our picture, rather, because the child's operations with sentences have fitted into his life in such a way as to permit us to suppose that the sentences have been serving as vehicles for the transmission of knowledge.

(Anyone who is inclined to suppose that this cannot be right, and that possession of content must be grounded in availability for belief-expression, needs to resist the temptation to rely on the dubious thought mentioned in §7: the thought that the possession of content is too mysterious to be understood except in terms of the peculiar properties of an ethereal medium, conceived as the stuff of which mental states are composed.)

9. The objects to which a formal semantic theory assigns semantic properties tend naturally to be rather formal objects: objects arrived at, then, by abstracting ruthlessly from the concrete detail of actual behaviour. In an access of mathematical enthusiasm about these formal objects and their formally assigned properties, it is possible to forget, or even deny, the anchoring of the discipline in the realities of speech or sufficiently speech-like behaviour. But semantics, formal or otherwise, is secondary to the theory of communication in at least this sense: if there were no subject matter for the theory of communication, there would be nothing for semantic theories to be about. Strawson goes further than this. He endorses a specific proposal about the nature of communication which I have found questionable. And in keeping with this view of communication, he takes the dependence of semantics on the

<sup>20</sup> 'Light dawns gradually over the whole': Wittgenstein, *op. cit.*, §141.

theory of communication to involve a possibility of analysing, in purely psychological terms, the properties which fit words and sentences to be objects of semantic theories; whereas I have suggested that such reductive analysis is not compulsory, and that it misleads about what must be understood in terms of what. Strawson's main aim, however, in the lecture I have discussed, is to recall formal semantics from platonistic excess, and remind it that a semantic theory is nothing if not a component in an account of actual or possible communicative behaviour. That he is correct in this is the unquestioned foundation of the suggestions, divergent in detail from the position he recommends, which I have made in this paper.

# The Individuation of Proper Names

L. Jonathan Cohen

Some of Sir Peter Strawson's most interesting work has been in connection with the ways in which certain features of natural language differ from the representations of them that are prevalent among formal logicians. The following is a latter-day effort that tends in a similar direction.

There are two different conceptions of a proper name. According to one, which I shall call the 'linguistic' conception, proper names are, and ought to be, individuated by their phonological form. Thus *Aristotle* is a different name from *Plato*, but the given name of the philosopher who tutored Alexander the Great is quite properly, and without any taint of equivocation, the same as that of the shipowner who married President Kennedy's widow. According to the other conception, which I shall call the 'idiosyncratic' one, proper names are, or would ideally be, individuated by their form-and-bearer. On this view it is either said that the philosopher's given name is different from that of the shipowner or, alternatively, that the form *Aristotle* is unfortunately at least two ways homonymous. The present paper is concerned to strengthen<sup>1</sup> the case for adopting the linguistic rather than the idiosyncratic conception, and to articulate some of the implications of so doing.

The idiosyncratic conception has been adopted by many modern logicians. Frege, like Russell, thought not only that in a logically perfect language every proper name would designate just one object, but also, presumably, that where a natural

<sup>1</sup> Some arguments in its favour were sketched in L. Jonathan Cohen, 'Searle's Theory of Speech Acts', *Philosophical Review*, lxxix, 1970, pp. 547-50. A further development of it was given by Tyler Burge, 'Reference and Proper Names', *Journal of Philosophy*, lxx, 1973, pp. 425-39 and by J. J. Katz, 'A Proper Theory of Names', *Philosophical Studies*, xxxi, 1977, pp. 1-80; but both Burge's and Katz's proposals are unsatisfactory in certain respects, which are discussed below.

language proper name designated now one, now another, object this was an imperfection and a kind of homonymy.<sup>2</sup> Quine wanted to replace *Socrates* by 'a general term that is true of just one object'.<sup>3</sup> Though Dummett admits that the doctrine 'requires qualification in many respects before it becomes a realistic picture of our actual employment of names', he thinks that in principle 'a proper name possesses a unique and specific sense, common to all users of the name, which determines its reference'.<sup>4</sup> According to Kripke the main problem about proper names is whether such a name determines its characteristically unique reference by virtue of its sense, as Frege thought, or in some other way.<sup>5</sup> According to David Lewis 'we would distinguish the name of London (England) from the homonymous name of London (Ontario)'.<sup>6</sup> And within the framework of the idiosyncratic concept an immense amount of controversy has been generated by the question whether a natural-language proper name has a sense as well as a designation. On the one hand the assignment of a sense, or even of a disjunction of senses, seems to restrict the range of possibilities that can be considered. If *Aristotle* means the philosopher who tutored Alexander the Great, we cannot self-consistently consider what might have happened if Aristotle had not performed that task. On the other hand, if a proper name has no sense, it seems to have nothing with which to pick out a unique designation, in the way that the meaning of a common name like *cat* picks out the manifold extension of that name. So causal theories are constructed which tend, as soon as their details are made sufficiently determinate, to be too easily confounded by counter-examples.

The linguistic conception side-steps this tiresome controversy altogether, and guides us towards a less problematic semantical characterization. At least three considerations

<sup>2</sup> Cf. 'On Sense and Reference', *passim*, in *Translations from the Writings of Gottlob Frege*, ed. P. Geach and M. Black, esp. p. 70.

<sup>3</sup> W. V. O. Quine, *Word and Object*, 1960, p. 179.

<sup>4</sup> M. Dummett, *Frege Philosophy of Language*, 1973, p. 142. Cf. also 'What is a Theory of Meaning?' in *Mind and Language*, ed. S. Guttenplan, 1975, p. 132, where Dummett's argument against Davidson presupposes the validity of the idiosyncratic conception of proper names.

<sup>5</sup> S. A. Kripke, 'Naming and Necessity', in *Semantics of Natural Language*, ed. D. Davidson and G. Harman, 1972, pp. 252 ff.

<sup>6</sup> David Lewis, 'Truth in Fiction', *Amer. Philos. Quart.* V, 1978, p. 39.

stand in its favour. The first two are easily stated: the third requires greater elaboration.

## II

First, it is of fundamental importance to grasp that the properness of proper names is a feature—in Saussurean terms—of ‘parole’, not of ‘langue’. In a particular utterance or sequence of utterances the name *Smith*, when functioning as a singular term, is typically peculiar to the person it names: it designates just one person. If more people of that name are at issue and the name is used without a distinguishing verbal or non-verbal context, then the use is ambiguous and deviant. But within any given natural language each proper name is normally available, on appropriately different occasions, for naming any number of entities of the appropriate category. During the course of one day a single speaker of English could unambiguously designate six different people by the name *Smith* if he managed to have six conversations that differed appropriately in their circumstances.

Perhaps someone will object that we should think of each such conversation as taking place in an appropriate sub-language of English, so that *Smith* can be conceived to have a unique designation in each sub-language. But this objection cannot be sustained. No doubt it seems flattering: it makes brilliant linguists of us all. But the flattery is empty, because for many other theoretical purposes, whether synchronic or diachronic, we shall want a less extravagant method of individuating languages. A theory of language in such a Pickwickian sense of *language* would have rather poor explanatory value. It is methodologically unrewarding to have to sacrifice the unity of the larger entity (a language) in order to preserve that of the smaller one (a proper name’s designation). Moreover, if you speak one language when you refer to your neighbour by *Smith* and a second when you refer to your colleague by that name, then what language are you speaking when they both come to your party and you proudly announce

There are two *Smiths* here this evening,  
or when, as does happen sometimes, you find occasion to refer separately to both of them in the same conversation? Someone

may perhaps speculate that in the small groups in which he supposes language to have originated the philological ancestors of modern proper names had unique designations. But our task is to describe language as we know it is now, not as we speculate it to have been in the remote and unknowable past.

Convenience, not logical principle, is the main factor determining how many objects bear the same name. For example, the larger and better known the object, the more important that its name should not be widely shared. Only a few towns in the world are called London, but many English villages are called Coombe, and even more Englishmen are called John. Occasionally special scruples prevent reuse of a name, as in the case of Japanese emperors' names, but such scruples are religious or ethical rather than linguistic. It is misleading to describe this versatility of proper names as a form of homonymy, since all proper names are intrinsically versatile whereas homonymy, as in *bank* or *bat*, is an inconvenience from which almost all words of a particular grammatical category—almost all common-noun-forms, almost all verb-forms, almost all adjective-forms, etc.—are normally free. Admittedly, languages sometimes extend the meanings of common nouns, verbs or adjectives, so that *blue* comes to signify a certain state of mind or a certain type of film, for example, as well as a certain colour. But this is an event in the history of the 'langue'—not an event in the history of 'parole', like giving an old name to a new baby or a new ship: it is characteristically recorded by dictionaries, not by newspapers. The idiosyncratic conception of a proper name therefore carries with it the suggestion, as in Frege's writings, that natural-language proper names have drifted away from univocation, and deviated into homonymy, by some series of philological accidents or through the ignorance, laziness, carelessness, or improvidence of those who baptize children, name towns, etc. Yet in fact it is an immense convenience to have a reusable reservoir of phonologically acceptable speech-forms, divided perhaps into several categories, which parents of children, founders of cities or discoverers of mountains can draw on when they need.

Of course, we must distinguish between simple proper names and compound ones. The simple proper names of a language—sometimes called its proper-nouns—are listed in



dictionaries of male given names, female given names, surnames, field-names, place-names, etc. for that language. They are a species of word, co-ordinate with common nouns, verbs and the other traditional parts of speech, in any language in which these distinctions are syntactically important. On the other hand, the compound proper names of a language, which should perhaps be called its proper noun-phrases, are constructed out of its proper nouns or other words according to a variety of patterns. Thus we can have in English such compounds as *The Open University*, *Mount Everest*, *The New York Guaranty Trust Company* and *Charles Algernon Montgomery Smith*. But even compound proper names are reusable. Though it may be illegal for you to name your little school *The Open University*, you have not committed any linguistic solecism. Similarly, if Smith's parents failed to give him a sufficient number of personal names to distinguish him from all past, present, and future bearers of the same surname, they may have been deficient in parental care, but scarcely in linguistic propriety. And, after all, Smith's date and place of birth can always eke out identification when this does need to be unique, as on a passport.

Thus it is far from being the case, as the idiosyncratic conception implies, that a natural-language proper name normally has a unique designation, or that in an ideal natural language it would have one. Just the opposite. It would be a real imperfection in a natural language to prohibit the reuse of any proper name. Such a prohibition might be workable in a small, static community over a short period. But the phonological inventiveness it would require within any other kind of speech-community would be quite beyond normal human powers. Just as language economizes a speaker's effort in saying different things by allowing an infinite variety of sentences to be composed out of a finite number of words, so too it economizes his effort in naming different objects by allowing an infinite number of objects to be designatable (though not all at once) with a finite set of proper names. The idiosyncratic conception of proper names obstructs awareness of this immensely convenient feature of human language. If a logician's linguistic intuition generates that conception, rather than the linguistic one, he would do well to disregard his intuition.

The second consideration that stands in favour of the linguistic conception of proper names is the nature of human knowledge about them. That the philosopher who tutored Alexander was called Aristotle is normally accepted as being a matter of historical fact, not of language. People who do not know that philosopher's name, when asked, are taken to show historical ignorance, not linguistic imperfection, and the same is true if they do not know the name of a living person who is pointed out to them. Learning the given names of our new neighbours, or even their full names, does not necessarily add anything to our knowledge of English: we might well have known already that the names *John* and *Mary*—which our new neighbours happen to have—are English names. But when we learn our new neighbours' names we have at least then learned a fact that we need to know in order to avoid excessive formality in addressing them.

It is therefore important not to be misled by the partial parallelism between the function of a proper name in natural language and that of an individual constant in an interpreted formal system. If you don't know what the individual constant designates, you haven't got a complete knowledge of this artificial language. But you can have a complete knowledge of English, including even its whole present-day vocabulary of proper names, and still not know what given name or what surname—let alone what full name—any particular person has. Certainly there is a sense in which a particular individual constant in an interpreted formal system could have had a different designation. But to have given it that other designation would have constituted an alteration in the artificial language-system; so, relative to the original system, the constant could not have had a different designation from the one it in fact had. On the other hand our new neighbour's parents could perfectly well have decided not to call their son *John*, or could have changed the family surname, without changing any feature whatever of their natural language.<sup>7</sup> The name that he has is as much a contingent fact about him (changeable by deed

<sup>7</sup> This is implicitly denied by S. Kripke, 'Identity and Necessity', in M. K. Munitz, *Identity and Individuation*, p. 155, where he assumes that a renaming of planets entails a change of language.

poll) as the colour of his car (changeable by respray). Our language would be volatile indeed if every time anyone got married and changed his or her name they thereby changed the language! Versatility, not volatility, is the hallmark of natural language.

It follows that where a mountain, a person or any other object has been given two different names, or where a person or corporation has adopted an alias, a statement asserting the identity of what the one name designates with what the other does is typically a contingent truth, not a necessary one. To describe it as a necessary, though a posteriori truth, would be to treat proper names in a natural language as if they were individual constants in an unalterable artificial language-system. When we learn the two names we learn two facts about an object which need not stand in any form of necessary connection with one another. Of course, some cultures might have elaborate rules for naming children that generated a form of necessary truth for the utterance of an appropriate identity-sentence in an appropriate context. Maybe in some remote Pacific island, peopled by English castaways, social conventions require that a man's name is John if and only if he is also called George. Once we know the rule we know, in a particular case, that John must be George. But the conceivability of such a bizarre culture serves merely to highlight the contingency of ordinary name-name identifications in natural language. If the ordinary ones were not contingent there would be no point in saying, on the Pacific island, that John must be George. More familiarly, perhaps, in the ancient Roman Republic anyone who bore the 'cognomen' *Cicero* was necessarily also named *Tullius*, since the Cicero family was part of the Tullian gens. But by these standards it is quite contingent for someone named *Nixon* to be also called *Richard*.

Philosophers who deny the contingency of any name-name identifications have tended to concentrate their attention on notoriously true identifications. Their thesis is intuitively less plausible when applied to false ones, like

(1) Bacon is identical with Shakespeare.

It scarcely seems impossible to imagine a course of events by which Nicholas Bacon's youngest son was born at Stratford and

wrote the plays under a pseudonym as well as the *Novum Organum* under his own name. But these events never actually occurred, so (1) is contingently false. We therefore cannot say that if (1) were true it would be necessarily true: if (1) were true it would be contingently true. Equally, denials of true identity statements often seem eminently plausible candidates for contingent falsehood. If I say

(2) Hesperus is not identical with Phosphorus,

I may well be imagining a possible system of planetary revolutions that allows for this non-identity. Perhaps you wrote off your natural inclination to take

(3) Hesperus is identical with Phosphorus

as a contingent truth, on the ground that the apparent contingency of (3) is due to its announcement of an empirical discovery. But (2) announces no such discovery. Its assertion could well be a piece of unempirical dogmatism. So the natural inclination to take (2) as a contingent falsehood cannot so easily be written off; and if (2) is contingent so is (3).

It follows that there is bound to be a fallacy somewhere in any argument designed to show that all true name-name identifications are necessarily true. One such argument, for example,<sup>8</sup> proceeds from the premiss that everything has the property of being necessarily identical with itself. Hence George has the property of being necessarily identical with George. And if John is identical with George, John must have every property that George has. Therefore, it is claimed, John has the property of being necessarily identical with George. But the argument is essentially question-begging in that it assumes the sharing of properties, which goes along with identity, to include properties involving modalities. Or, alternatively, when the argument is set out formally, we can say that the fallacy consists in allowing modal operators to appear within substituends for predicate variables. Not that this procedure ought always to be formally impermissible. But it should

<sup>8</sup> Cf. S. Kripke, *op. cit.*, pp. 135 ff.

not occur in a formal system that is designed to represent the logical structure of natural language.<sup>9</sup> Each syntactically differentiated kind of symbol-occurrence in such a formal system should be reserved, in the intended interpretation, for the performance of just one linguistic function, and just one kind of symbol-occurrence for each function: predicates for predicates, operators for operators, and so on. Go against this principle in your formalizations, if you please; but you are not then entitled to use your consequentially distorted representation of natural language as a basis for drawing related conclusions about what natural language is really like, as if you had not introduced any relevant distortions by your very system for representing its structure. After all, you would not judge your own appearance from its image in a distorting mirror.

Perhaps it will be objected that *necessarily* can occur within substituends for natural-language predicate variables if it occurs as a predicate modifier to express modality *de re*, and not as a sentential operator to express modality *de dicto*, and that its availability for use as such a predicate modifier is supported by the semantics for that use which has now been constructed.<sup>10</sup> But, in order to show that *necessarily* actually functions as a predicate modifier in relevant contexts in natural language, we need rather better reasons than the existence of this semantics. We need to be shown not merely what the semantics of such a modal predicate-modifier would be if the modifier were part of natural language. We need to be shown also that it is a relevant part of natural language. We need to be told of a statement like

Hesperus is necessarily identical with Phosphorus

that would have to be recognized both as true, and as having a thus modified predicate, even by those who adopt a semantics for proper names like the one detailed here: otherwise the objection would be question-begging. And the anti-essentialist has in effect always denied that any such statements exist.

\* The fallacy here is discussed at greater length, and in relation to other examples, in L. Jonathan Cohen, *The Implications of Induction*, 1970, pp. 211 ff.

<sup>10</sup> D. Wiggins, 'The *De Re* "Must": A Note on the Logical Form of Essentialist Claims' in *Truth and Meaning*, ed. G. Evans and J. McDowell, 1976, pp. 285 ff., and C. Peacocke, 'An Appendix to David Wiggins' "Note", *ibid.*, pp. 313 ff.

IV

The third consideration in favour of the linguistic, rather than the idiosyncratic, concept of proper names is their actual mode of linguistic operation. In order to understand how proper names normally function in a natural language one has to grasp that this is a form of indexical, or token-reflexive usage. Logicians are ready enough to grant that demonstrative pronouns (like *he*, *outside*, *to-day*, *hereafter*), temporally tensed verbs (like *jumps* and *jumped*), or temporally oriented adjectives (like *present* and *future*) can have their designations determined, in actual speech, by their context of utterance. But they have generally been reluctant to recognize that this is how proper names normally function also.

Admittedly Tyler Burge<sup>11</sup> has recently argued that proper names function both as demonstratives (in *Alfred studies in Princeton*) and as predicates or general terms (in *There are relatively few Alfreds in Princeton*). But even this does not get it quite right. The concept of a demonstrative is a fairly clear-cut one, which loses a lot of its value if we blur it in order to include proper names within its extension. Demonstratives have their designations determined by their context of utterance without the benefit of any previous baptismal ceremonies, official registrations, etc. They are extremely adaptable tools of reference which a speaker can use without having any previous knowledge whatever about the entity designated (except perhaps knowledge of some highly general feature such as its sex). For example, the demonstrative *I* can be used in this way, without any impropriety or risk of misunderstanding, by a known amnesiac. But demonstratives pay the price for their extreme adaptability by being rather more difficult to use when the speaker is at a distance—in time, space, or subject-matter—from what he wishes to designate. Demonstratives are typically helped to achieve their designation by their relative proximity, along one dimension or another, to what they designate: the easiest things to direct attention to, for purposes of

<sup>11</sup> Opponent Burge's theory is endorsed by J. Hornsby, 'Singular Terms in Contexts of Propositional Attitude', *Mind*, lxxxv, 1977, p. 42, and criticized (from another point of view than mine) by S. E. Boer, 'Proper Names as Predicates', *Philosophical Studies*, 27,

mutual understanding, are what you and your addressee can both see or have just been discussing.<sup>12</sup> The use of a proper name, on the other hand, does not depend in any way on proximity to what it designates. No doubt proper names are substantially less flexible than demonstratives, and their optimally effective use requires both the speaker and his addressee to have stored in their memory certain appropriate items of non-linguistic knowledge. But proper names have the counter-vailing advantage that they can be used to designate entities which are far beyond the speaker's current range of pointing out. No useful purpose, therefore, can be served by confounding together two such valuably different categories of words. Even if we were to classify proper names as a species of 'demonstrative', in some extended sense of that word, we should still need another term to cover those 'demonstratives' that are not proper names.

Nor is it particularly significant, as Burge seems to imply, that proper names sometimes function as predicates or general terms, in sentences like

There are relatively few Alfreds in Princeton.  
Burge objects to treating such uses as special ones, and wishes to regard modified and unmodified occurrences of proper names as co-ordinate types of usage. But Burge offers no evidence that the modified use of proper names occurs in most natural languages: all his examples are from English. And the fact is that all forms of indexical expression tend to lend themselves, at least in some languages, to de-indexicalization. This certainly happens with ordinary English demonstratives, though Burge seems unaware of it.<sup>13</sup> Compare, for example,

- (4) That cat is a *he*.<sup>14</sup>

(6) To-day's problem is always more exciting than yesterday's

or

(7) Some people fear the *hereafter*.

And it also happens with temporally tensed verbs, as in

(8) Kittens *jump* when startled,

or with temporally oriented adjectives, as in

(9) *Present* love has *present* laughter.

In sum, ability to function as a descriptive term has not been established as a feature of proper names that distinguishes them from other indexically usable expressions;<sup>15</sup> and a technique for formalizing sentences that include proper names is obliged to represent a connection between their indexical and descriptive uses only so far as it also needs to represent a connection between sentences like

That cat is a he

and

He is purring.

Indeed, so far as it is the task of logical formalizations to represent such connections, one might think it more important to connect the designatory potential of proper names with their vocative use. Proper names are very commonly used to request or arouse the attention of the persons to whom they can, in those contexts, also be used to refer. Burge does not mention this use, though it is much more closely confined to proper names, among indexically usable expressions, than is deindexicalization: perhaps the second person pronoun is the only demonstrative that regularly functions thus. Yet any formalization of imperative sentences like

John, come here!

must be capable of representing vocatives.

<sup>15</sup> Not that there are not other kinds of English word of which it seems more appropriate to say, as Burge says of proper names, that usage A is just an indexicalization of usage B, rather than that usage B is just a deindexicalization of usage A. The names for days of the week are a typical example, as in (usage A) *He will arrive on Sunday*, as distinct from (usage B) *Sundays are sacred*.



How then does the context of a proper name's utterance determine its designation? We need to distinguish at the outset here between the object that a speaker intends to designate by his utterance of a proper name, the object that is actually designated, and the object that the addressee takes to be designated.

By a slip of the tongue, or as a result of originally mishearing, or later misremembering, the Peripatetic philosopher's name you might utter the sentence

(10) Aristophanes tutored a great conqueror.

You have it clearly in mind that Alexander was tutored by a philosopher, about whom you know a number of facts such as that he was born at Stagirus in 384 BC and walked about while lecturing; and you certainly do not think that Alexander was tutored by a very aged comic playwright. Whom does the name *Aristophanes* in your utterance designate? According to many logicians<sup>16</sup> it designates the person you had in mind or meant, viz. Aristotle. But 'designates' here is in any case a technical term and we have to adopt those criteria for its predication that are appropriate to the theoretical function of statements about designations. And these statements must at least give the facts that are needed for truth-assessments, if statements about truth-conditions form at least a part of any adequate semantics for a natural language. Hence 'designates' must be so defined that what *Aristophanes* designates in your utterance of (10) must, if and only if your utterance is true, be an entity satisfying the predicate *tutored a great conqueror*. So we need to decide whether to call your utterance of (10) true, or false. Well, what did you say? You said that Aristophanes tutored a great conqueror. You may have intended to say that Aristotle did this, but what you actually said was that Aristophanes did it. No court of law, no jury of your colleagues, no verdict of public opinion, would describe what you said otherwise. Reformulation of your utterance in indirect speech supplies a decisive basis for truth-assessment. After all this is the basis on which the biconditionals of truth-conditional semantics are constructed. The sentence 'Snow is white' is held to be true, for any speaker at any

<sup>16</sup> e.g. M. Devitt, 'Singular Terms', *Journal of philosophy*, lxxi, 1974, p. 488.

time, if and only if snow is white, just because, whatever such a speaker intends to say, what he actually says is that snow is white. It follows that, irrespective of your intentions or of your beliefs or of what you had in mind or of the causal origins of your mistake, your utterance of (10) was false. The name *Aristophanes* in your utterance did not designate the man whom you had in mind or meant. It is just the same with demonstrative pronouns. Suppose you say

(11) He is coming with me,

having John in mind and pointing over your shoulder, since you think that John is behind you; and suppose that Jim, not John, is the man behind you. Despite what you intended, what you have actually said is that Jim is coming with you and the demonstrative pronoun *he* in your utterance of (11) designated Jim, not John.<sup>17</sup> The objective designation of your pronoun depends not on what you yourself happened to believe, but on what an observer of your gesture might reasonably presume that you knew.

No doubt there is a sense in which you were referring to Aristotle, when you unfortunately used the wrong name for him, and to John; when you pointed in the wrong direction. But, even though it would be quite true, in one ordinary, everyday sense of *refer* to say this, that sense of *refer* is not the one needed to do the technical job for which I have been using the word *designate*. In the pragmatic, or speech-act, sense of *refer*, people and their utterances refer, and frequently what they refer to doesn't even exist, though they may think it does, like the book which Professor X referred to in his lectures and which turns out never to have been written. In the usual semantic, or logical, sense of *refer*, however—i.e. in the sense expressed by *designate* in the present paper—it is only words, phrases, individual constants, class abstracts, etc. that refer, and existence is arguably a necessary condition for being referred to.<sup>18</sup> Again, in the everyday, pragmatic sense of *refer* you have not referred to John, but addressed him, when you say

<sup>17</sup> J. J. Katz, 'A Proper Theory of Names', *Philosophical Studies*, xxxi, 1977, pp. 1-80, has a linguistic token-reflexive theory of proper names, but spoils it by identifying the actual reference of a proper name's utterance with its intended reference.

<sup>18</sup> On the difference between these two senses of *refer*, see further L. Jonathan Cohen, 'What is the ability to refer to things, as a constituent of a language-speaker's competence?', in R. Jakobson et al., *Linguaggi nella Società e nella Tecnica*, 1970, pp. 259 ff.

John, come here!

Yet this need not prevent our saying that the first word of your utterance designates the subject of its imperative verb-form.

Very similar considerations refute other psychologizing accounts. For example, even though the person with whom you are talking is familiar with your tendency to say 'Aristophanes' when you mean 'Aristotle', it does not follow from that fact that the proper-name token in your utterance of (1) actually designates Aristotle. We have to distinguish here between the linguistic meaning of your utterance and the message that it successfully communicated.

How then is it determined what such or such an occurrence of a particular proper name actually designates? Four factors seem capable of playing a critical part, though they are not equally important in every utterance: the topic under discussion, the speaker's knowledge about what entities within that topic have the name in question, his addressee's knowledge about what entities within the topic have the name in question, and any relevant events, circumstances, etc. that are present at the time and place of the utterance and may be reasonably presumed to be known to the speaker and his addressee. In a book, for example, where the author of the text is rather remote from his readers and the readers may be rather heterogeneous in their education, the designation of a proper name will be primarily determined by the book's topic. In a history of logic *Aristotle* denotes the Stagirite; in memoirs of the twentieth century jet set it denotes the shipowner. In the course of a brief conversation, however, the topic of a remark may be insufficiently clear to fix the designation of a particular proper name, and the determining consideration may be the fact that only one person of that name is known to both the speaker and his addressee. Nor need they both have precisely the same knowledge. Maybe the speaker knows only that the founder of the Peripatetic school of philosophy had the name *Aristotle*, while his addressee knows only that the most famous person to be born in Stagirus had that name. But in virtue of knowing these two things they both know enough to establish the name's designation. Or perhaps the speaker and his addressee share knowledge of several people with the name in question, and the

about the fact of name-acquisition, but neither need know that the other knows. A bystander could always determine the designation by looking only at the four factors already listed.

Finally we shall perhaps be told that, if the speaker wishes to use a proper name in a successful linguistic communication, he must know that the addressee knows the name's bearer, and that, if the addressee is to interpret the speaker's utterance correctly, he must know that the speaker knows the name's bearer: so in general there must be common knowledge that members of the community have in their repertoire the procedure of using the name to refer to the object. But one needs to distinguish here between the conditions requisite for a successful act of communication in which a proper name is employed, and the conditions determining what the proper name actually designates. It is just because the latter may include the first-order knowledge-states of speaker and addressee that the former may have to include knowledge about these knowledge-states. To know what the name's occurrence designates, one must know what relevant conditions are present.

VI

The idiosyncratic conception of proper names, therefore, with its projected individuation of them by form-and-bearer, gains no support from considerations about the way in which natural language actually functions. Nevertheless the requirements of theoretical systematization often force us to oversimplify. Ideal gases, lines that have length without breadth, and perhaps also truth-functional connectives, are familiar examples of such well-motivated regimentation. They constitute misrepresentations only so far as we have some good reason to conceive our theories as descriptions, rather than idealizations, of their several domains. So logicians might be tempted to defend an idiosyncratic conception of proper names on the ground that it allows a very simple and straightforward procedure for the formalization of sentences containing proper names. Each proper name, when individuated by form-and-bearer, is assumed to correlate with a distinct individual constant in the formalism.

Now there is certainly no logical harm in this mode of formalization. It will not lead us from true premisses to false conclu-

sions; and, given a sufficiently inexhaustible supply of individual constants, our formal language can represent thus (other things being equal) any argument that is constructable in natural language with the help of proper names or demonstratives. Indeed this mode of formalization is ideal for logical exercises and for the analysis of particular pieces of actual reasoning. But it will not serve as a satisfactory representation of how proper names function in natural language because, when operating as such a representation, it individuates proper names quite incorrectly and does not reveal the essentially indexical, or token-reflexive way in which their designations are determined. Nor, correspondingly, should this mode of formalization be used in the construction of a truth-conditional semantics for a natural language, if the theorems of that semantics are to state truth-conditions for sentences of the language that are identified and individuated by their appropriate structural description.

One method of dealing with proper names in a truth-conditional semantics, on a non-idiosyncratic basis, has been proposed by Burge.<sup>22</sup> It exploits the point that all English proper names have a descriptive or predicative use. His analysis for a sentence like (12) is a formal paraphrase of the sentence

For all  $x$  and  $y$ , if  $x$  is an act of reference by a speaker  $s$  at time  $t$  to  $y$  with the first occurrence of *Aristotle* in *Aristotle is a philosopher*, then *Aristotle is a philosopher* is true with respect to  $s$  at  $t$  just in case the object which is  $y$  and is an Aristotle is a philosopher.

However, not only are Burge's proposals of dubious value for languages other than English. They also depend essentially on the mistaken assumption that what is relevant to the truth of an utterance is the speaker's subjective act of referring rather than the objective designation that a normal bystander would attribute to the utterance. Only on the former assumption would it be reasonable to restrict the relativization of a proper name's

<sup>22</sup> 'Reference and Proper Names', *Journal of Philosophy*, lxx, 1973, p. 433. I am accepting for present purposes the general framework of truth-conditional semantics, as a way of talking about meanings in natural language, though in a more extended discussion this acceptance would have to be subject to some important qualifications.

titles in much the same way as it relativizes those of proper names.

What should we say about Pegasus, Pickwick, 'et hoc genus omne'? Proper names may be given not only to actual persons, animals, places and events, but also to deities, angels and devils, to mythological heroes, to the persons and places of creative fiction, to military operations that may well be cancelled before they take place, or even to numbers that may one day be proved not to exist (like the lowest even number that is not the sum of two primes, which we might name *Eldorado*).

It is sometimes suggested that the problem of fictional reference may be solved by assuming that all references in works of fiction are pretended ones. What we should have to consider, on this view, are not some special, fictional truth conditions for the sentence *Spenlow came out of the house* that might be used to make the fictional reference to a Mr Spenlow, but just ordinary truth conditions for a sentence asserting Charles Dicken's pretence that Spenlow came out of the house. But this suggestion will hardly work. Very few of the problematic references are pretended ones. Even in novels the author very rarely pretends to be referring to real people. To make such a pretence seriously, he needs to cite alleged sources or print photographs of mimeographed police dossiers, and so on. Most modern novels instead print an explicit disclaimer of any intention to refer to a real person, living or dead, so as to avoid the consequences of possible libel actions. One does not want to lose the distinction between fiction which pretends to be history and fiction which does not (or even denies that it is). In any case the generals who name an operation that they later have to cancel are not necessarily pretending anything when they refer to the operation by name in discussing its strategy and logistics: planning does not necessarily entail make-believe.

Indeed, one really should not be concerned, in an analysis of natural language, with the different domains of discourse to which speakers of the language devote their attention on different occasions. It is no business of linguistics to give some kind of priority to the physical over the spiritual, the historical over the fictitious, the achieved over the merely intended, or the confirmed over the merely hypothesized. Metaphysics should not be allowed to intrude into the semantics of natural

language. Language is an all-purpose instrument and allows references to be made, in appropriate contexts, to the elements of any conceivable domain. So all that needs to be said about (13) in this connection is that the term 'entity' in it should be taken to permit a corresponding latitude. Since one of the four determining factors for a proper name's designation is the subject-matter of the utterance in which it occurs, we are entitled to suppose that the entity designated merely has to fall within the utterance's domain of discourse. Of course, occasions occur when even this requirement is not satisfied. Genuine mistakes are sometimes made as to whether a particular domain of discourse contains an entity bearing such-or-such a proper name. But it is of no relevance to the present enquiry whether one should adopt a Russellian or a quasi-Strawsonian method of dealing with those eventualities. In (13) truth-conditions are given only for cases where the existential condition is satisfied; and acceptance of (13) leaves it an open question whether utterance of (12) should be regarded as an assertion of a false saying in the other cases, where the existential condition is not satisfied, or as an assertion of a saying with a neutral truth-value.

One final qualification needs to be added, however. Some languages divide their proper names into morphologically distinct categories, whereby people and places, for example, or men and women, can never bear the same names. Where this happens, the proper names have a good claim to be attributed corresponding meanings.<sup>24</sup> Thus in Latin the names *Julia* and *Julius* might be said to differ in meaning because their bearers must always differ in sex. But the point is easily represented within an analysis like (13). If English is such a language and *Aristotle* a male person's name within it, all that needs to be done to (13) is to replace 'entity' by 'male person'.

<sup>24</sup> Exact translation of a proper name, from one language to another, may thus be rendered rather difficult. But in any case it is only by a rather crude convention that we assume the Stagirite to have been given the name *Aristotle* rather than its Greek original, which was a different speech-form and therefore a different name. The crudity of this convention generates serious problems for the analysis of indirect discourse, which are interestingly discussed in Saul Kripke, 'A Puzzle about Belief', in *Meaning and Use*, ed. A. Margalit (1979), pp. 239-83.

predicate to the name in question, others of them treating of quite other matters. Now this whole conglomerate sentence can be transformed into a single predication, having the name in question as its subject and a relative clause as predicate—a conglomerate relative clause, embodying the full complexity of the original sentence. The clause is formed from the original or source sentence by supplanting the name in question by a relative pronoun and its recurrences by an ordinary pronoun and adjusting the word order. Everything that any sentence says about something can be packaged thus in a single complex predicate; such is the relative clause.

Grammatically it may be viewed as an adjective phrase.<sup>2</sup> Strictly speaking it is used only attributively, rather than predicatively, and is preceded by the substantive that it governs. But we can form a substantive from it by prefixing a dummy substantive ('thing', 'person') and then we can predicate the whole with help of 'is a'. Not that we want to; the predication would simply be equivalent to a source sentence of the relative clause, thus getting us nowhere. The reason to contemplate predicating relative clauses is just in order to explain the relative clause, as I did, by saying that the predication is equivalent to the corresponding source sentence. The relative clause has its utility rather in attributive uses that do not resolve out thus through predication.

I said that the use of pronouns basic to our inquiry into reference, or ontology, is the use in relative clauses. Here the pronoun is the relative pronoun 'that' or 'which' (or 'who' or 'whom'), together with its recurrences in the guise 'it' (or 'he' or 'she' or 'him' or 'her'). A variant of the relative clause, favoured by mathematicians for the simplicity of its syntax, is the 'such that' construction; the form 'which ... it ...' gives way to 'such that ... it ... it ...'. This form spares us the contortions of word order that are called for in manœuvring the relative pronoun into an early position. Also it facilitates the introduction of variables in lieu of pronouns so as to keep cross-references

<sup>2</sup> In 'Quine's syntactical insights' (D. Davidson and J. Hintikka, eds, *Words and Objections*, Reidel, 1969, pp. 146–57) Peter Geach has ably argued that the relative clause in English is not adjectival, nor even a grammatically coherent whole. The anomalies he adduces may afford clues to the genesis of the construction. The utility of relative clauses is best understood, even so, by thinking of them as complex general terms, adjectival in form, derived from sentences.



under control in the complex cases so typical of mathematics. This use of variables, in lieu of the relative pronoun and its recurrences, is the basic use of bound variables.

If pronouns were used only in the fashion of what Geach distinguishes as pronouns of laziness, they would have no special bearing on questions of reference; they would stand as mere abbreviations for their grammatical antecedents, and there would be no presumption that those antecedents were designations. Bound variables, if used in lieu only of such pronouns, would be substitutional rather than referential; they would be presumed only to admit expressions as substitutes and not objects as values.

However, matters do not remain thus. The pronoun or variable soon goes unequivocally referential when the relative clause is used in uneliminable positions, positions where it cannot be resolved back into a source sentence through predication. This happens, notably, when the substantivized relative clause 'things which . . .' is used as the subject term of a categorical sentence—thus preceded by 'every' or 'some'. Here there is no predication for us to equate to a source sentence of the relative clause. Nor, usually, can we expand the whole categorical into some great conjunction or alternation of singular sentences; for usually we have no exhaustive fund of designations at our disposal for the several things fulfilling the relative clause.

The pronoun or variable admits substitutions of constant singular terms even where these do not exhaust the relevant objects of reference. Such terms, as far as they go, are then properly seen as designating objects. This is how the pronoun or variable is primary and the name parasitic. Were it not for the irreducibly referential pronoun, or some idiom to the same effect, any distinction between designative words and others would be idle and arbitrary. Words would still be learned by conditioning to stimulation; and sentences would still be true and false, but a notion of objects would have no place.

Given the irreducibly referential uses of pronouns or variables, on the other hand, we naturally bring the pronouns of laziness in tow and treat them as referential too, except in some cases where, by defining them away, we can lop off some unwelcome ontological excrescence. Thus the *elle* of the French